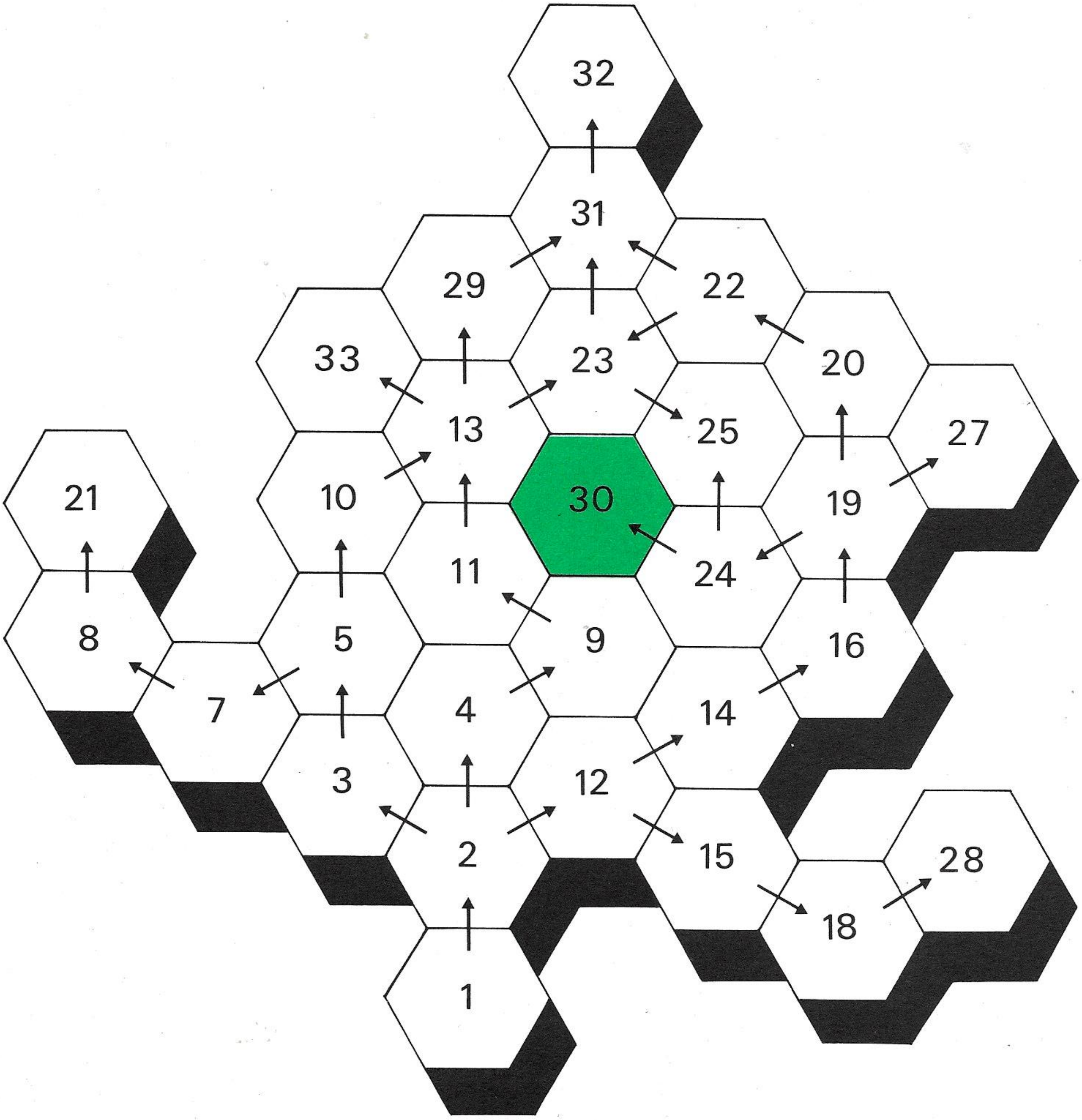
M201 30

THE OPEN UNIVERSITY

Mathematics: A Second Level Course

Linear Mathematics Unit 30

Numerical Solution of Eigenvalue Problems



The Open University

Mathematics: A Second Level Course

Linear Mathematics Unit 30

NUMERICAL SOLUTION OF EIGENVALUE PROBLEMS

Prepared by the Linear Mathematics Course Team

The Open University Press

The Open University Press Walton Hall Milton Keynes, MK7 6AA

First published 1972. Reprinted 1976

Designed by the Media Development Group of the Open University.

Printed in Great Britain by
Martin Cadbury

SBN 335 01120 9

This text forms part of the correspondence element of an Open University Second Level Course. The complete list of units in the course is given at the end of this text.

For general availability of supporting material referred to in this text, please write to the Director of Marketing, The Open University, P.O. Box 81, Walton Hall, Milton Keynes, MK7 6AT.

Further information on Open University courses may be obtained from the Admissions Office, The Open University, P.O. Box 48, Walton Hall, Milton Keynes, MK7 6AB.

1.2

Contents

Set Books

D. L. Kreider, R. G. Kuller, D. R. Ostberg and F. W. Perkins, *An Introduction to Linear Analysis* (Addison-Wesley, 1966).

E. D. Nering, *Linear Algebra and Matrix Theory* (John Wiley, 1970).

It is essential to have these books; the course is based on them and will not make sense without them.

Conventions

Before working through this correspondence text make sure you have read *A Guide to the Linear Mathematics Course*. Of the typographical conventions given in the Guide the following are the most important.

The set books are referred to as:
K for *An Introduction to Linear Analysis*
N for *Linear Algebra and Matrix Theory*

All starred items in the summaries are examinable.

References to the Open University Mathematics Foundation Course units (The Open University Press, 1971) take the form *Unit M100 3, Operations and Morphisms.*

Note

Please note that this text is not based on the set books for the course.

30.0 INTRODUCTION

The concepts of eigenvalue and eigenvector have played an important role in this course, particularly in *Unit 13, Systems of Differential Equations, Unit 23, The Wave Equation,* and *Unit 25, Boundary-value Problems.* We introduced these concepts in *Unit 5, Determinants and Eigenvalues.* If σ is a linear transformation from a vector space V to itself, the scalar λ is called an eigenvalue, and the vector ξ an eigenvector, if $\xi \neq 0$ and $\sigma\xi = \lambda\xi$. In this unit we shall assume that V is finite-dimensional and real (typically, R^n) so that this equation can be written as

$$A\mathbf{x} = \lambda\mathbf{x}$$

where A is a real $n \times n$ matrix and $\mathbf{x}$ is a one-column $n \times 1$ matrix. We shall discuss numerical methods for calculating the eigenvalues $\lambda_1, \lambda_2, \ldots$ and the column matrices $\mathbf{x}_1, \mathbf{x}_2, \ldots$ representing the corresponding eigenvectors, when the square matrix A is given.

Unlike the problems of our numerical work in earlier units, in which the requirements were few and obvious, here we may have various requirements which will influence our choice of method. In some particular context we may require just the eigenvalue of largest modulus or magnitude; in another we may seek the eigenvalue of smallest modulus or the one nearest to a particular real number, and in yet another we may want all the eigenvalues. In each of these cases we may or may not be interested in the corresponding eigenvectors, but in this unit we concentrate on methods that give the eigenvector as well as the eigenvalue; both together are called an *eigensolution.*

When we want just one or two eigensolutions we are likely to use an iterative method. If we want the *complete eigensystem* (all the eigensolutions) we shall preferably use a method involving *similarity transformations* of A to a simpler similar matrix $P^{-1}AP$ (sub-section 10.1.1 of *Unit 10, Jordan Normal Form*). The iterative methods are treated in Section 30.3 and the use of similarity transformations in Section 30.4.

Following our standard practice in numerical work we must pay close attention to possible instabilities: the inherent instability of the problem and the induced instabilities of our various methods. We treat inherent instability in Section 30.2, and we shall examine each possible induced instability whenever we introduce a new method or apply an old method. As so often happens in numerical work we shall find that some quite obvious and attractive methods can be completely useless for non-trivial problems! In particular, the method based on solving the characteristic equation

$$\det(A - \lambda I) = 0,$$

which we found useful in *Units 5, 10* and *13* for 2×2 and sometimes for 3×3 matrices, is extremely unstable when applied to larger matrices with numerical entries.

In *Unit 8, Numerical Solution of Simultaneous Algebraic Equations,* we considered possibilities of improving the accuracy of a computed solution by doing just a little more work. In the eigenvalue problem we can do the same sort of thing, not only improving an approximate solution but finding useful estimates for its error. The relevant analysis is treated in Section 30.5, which is optional.

We shall concentrate exclusively on real symmetric matrices, which occur very frequently in practical problems. The mathematical properties of the symmetric case are considerably easier than those of the general case, and as a corollary of this the computing techniques are easier to understand and to analyse. The treatment of unsymmetric matrices is rather more difficult, and the only comment we make here is that very few of the results and techniques for the symmetric case carry over unchanged to the general case. We shall also assume, though this is not a significant restriction in practical problems, that the number of distinct eigenvalues is equal to the order of the matrix; that is, all the eigenspaces have dimension 1.

In Section 30.1 we shall list those mathematical properties of the symmetric eigenvalue problem which we shall need for our numerical techniques. Some of these properties have been treated in earlier units.

30.1 MATHEMATICAL PROPERTIES OF THE SYMMETRIC EIGENVALUE PROBLEM

30.1.1 Preliminary Discussion and Results

(i) *What Vector Space to Use?*

In most of this course we have thought of matrices as the numerical representation of linear transformations of various finite-dimensional vector spaces. In this unit, however, we shall be concerned *only* with matrices, and to avoid having to invent transformations for them to represent, it is convenient to work in a vector space V whose elements are themselves matrices. We take as our vector space V the set of all one-column matrices with n entries, and denote its elements by letters such as $\mathbf{x}, \mathbf{y}, \ldots$. We shall also need a Euclidean structure on this space: it is given by the standard inner product

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + \cdots + x_n y_n$$

that is

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x} \qquad (1)$$

where $\mathbf{x}^T$ is the transpose of the matrix $\mathbf{x}$.

(ii) *The Eigenvector Basis*

Any $n \times n$ matrix A defines a linear transformation

$$A : \mathbf{x} \longmapsto A\mathbf{x}$$

in our vector space of $n \times 1$ column matrices. If A is symmetric we can apply Theorem 5 of *Unit 24*, *Orthogonal and Symmetric Transformations*. This tells us that V has an orthonormal basis consisting of eigenvectors of A. We shall call the vectors in this basis $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. The orthonormality condition by (1), is

$$\mathbf{x}_r \cdot \mathbf{x}_s = \mathbf{x}_r^T \mathbf{x}_s = \delta_{rs} \qquad (2)$$

where δ_{rs} means 1 if $r = s$ and 0 if $r \neq s$.

(iii) *The Modal Matrix*

The matrix whose columns are $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ is often called the *modal matrix*, because its columns often represent modes of vibration of some physical system (see *Unit 13*, *Systems of Differential Equations*, especially the television programme). We denote it by

$$X = \begin{bmatrix} x_{11} \ldots x_{1n} \\ \vdots \qquad \vdots \\ x_{n1} \ldots x_{nn} \end{bmatrix}$$

Because of (2) it satisfies

$$X^T X = I$$

and is therefore an orthogonal matrix (see Theorem 3 of *Unit 24*). As we have verified in sub-section 10.2.1 of *Unit 10*, X has the property

$$AX = \begin{bmatrix} \lambda_1 x_{11} \ldots \lambda_n x_{1n} \\ \vdots \qquad \vdots \\ \lambda_1 x_{n1} \ldots \lambda_n x_{nn} \end{bmatrix} = XD$$

that is

$$X^{-1} A X = D \qquad (3)$$

where

$$D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & \cdot \\ & & & \lambda_n \end{bmatrix}$$

Equation (3) can also be interpreted by saying that X is a matrix of transition that diagonalizes A.

(iv) *Coordinates with respect to an Eigenvector Basis*

Since $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is a basis, any vector $\mathbf{y}$ can be expressed in the form

$$\mathbf{y} = \alpha_1 \mathbf{x}_1 + \cdots + \alpha_n \mathbf{x}_n$$

with suitable real numbers $\alpha_1, \ldots, \alpha_n$, which are the coordinates of $\mathbf{y}$ with respect to this basis. Since the basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is orthonormal, we have

$$\mathbf{y} \cdot \mathbf{y} = \mathbf{y}^T \mathbf{y} = \alpha_1^2 + \cdots + \alpha_n^2$$

As we saw in *Unit 16, Euclidean Spaces I* (page K278) the coefficients α_r can be calculated by taking the inner product with $\mathbf{x}_r$ and using the orthonormality condition (2): this gives

$$\alpha_r = \mathbf{x}_r^T \mathbf{y} = \mathbf{y}^T \mathbf{x}_r \qquad (4)$$

(v) *Quadratic Forms and Eigenvalues*

In *Unit 14, Bilinear and Quadratic Forms,* we defined a matrix A to be positive definite if and only if the quadratic form $\mathbf{y}^T A \mathbf{y}$ is positive for all non-zero vectors $\mathbf{y}$. We can express the condition for A to be positive definite in terms of its eigenvalues, since

$$\begin{aligned} \mathbf{y}^T A \mathbf{y} &= \mathbf{y}^T A(\alpha_1 \mathbf{x}_1 + \cdots + \alpha_n \mathbf{x}_n) \\ &= \mathbf{y}^T(\alpha_1 \lambda_1 \mathbf{x}_1 + \cdots + \alpha_n \lambda_n \mathbf{x}_n) \\ &= \alpha_1^2 \lambda_1 + \cdots + \alpha_n^2 \lambda_n \end{aligned} \qquad (5)$$

(from (4)), where $\mathbf{x}_r$ is the eigenvector of A corresponding to the eigenvalue λ_r. From Equation (5) we see that A is positive definite if and only if all its eigenvalues are positive.

From Equation (5) we can also derive a lower bound on $|\lambda_{\max}|$ where $\lambda_{\max}$ is defined as the eigenvalue of largest magnitude. Equation (5) gives

$$\begin{aligned} |\mathbf{y}^T A \mathbf{y}| &\leqslant \alpha_1^2 |\lambda_1| + \cdots + \alpha_n^2 |\lambda_n| \\ &\leqslant (\alpha_1^2 + \cdots + \alpha_n^2)|\lambda_{\max}| \\ &= (\mathbf{y}^T \mathbf{y})|\lambda_{\max}|. \end{aligned}$$

In particular if $\mathbf{y}$ is normalized (has unit length) then we have

$$|\mathbf{y}^T A \mathbf{y}| \leqslant |\lambda_{\max}| \qquad (6)$$

(vi) *Measuring Matrices*

Let us denote by $M(\mathbf{x})$, where $\mathbf{x}$ is any one-column matrix, the magnitude of the entry in $\mathbf{x}$ with largest magnitude, so that for example

$$\mathrm{M}\left(\begin{bmatrix} 2 \\ -3 \end{bmatrix}\right) = 3.$$

Let us also denote by $M(A)$ the largest of the row sums of the magnitudes of the entries of A: for example, if

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 0 \\ 3 & 0 & -2 \end{bmatrix}$$

the row sums are

$$1 + 2 + 3 = 6,$$
$$2 + |-1| + 0 = 3,$$
$$3 + 0 + |-2| = 5$$

and so $M(A) = 6$. Then it is not hard to show that

$$M(A\mathbf{x}) \leqslant M(A)M(\mathbf{x}).$$

This is true, in particular, if $\mathbf{x}$ is an eigenvector $\mathbf{x}_m$ corresponding to $\lambda_{\max}$, in which case the left-hand side is equal to

$$M(\lambda_{\max}\,\mathbf{x}_m) = |\lambda_{\max}|M(\mathbf{x}_m)$$

We then have that

$$|\lambda_{\max}|M(\mathbf{x}_m) \leqslant M(A)M(\mathbf{x}_m)$$

which gives

$$|\lambda_{\max}| \leqslant M(A) \qquad (7)$$

By combining this upper bound with the lower bound (6) we obtain the useful inequality

$$|\mathbf{y}^T A\mathbf{y}| \leqslant M(A), \text{ if } \mathbf{y}^T\mathbf{y} = 1. \qquad (8)$$

We can also show, by an argument which we do not give in detail here (it uses the Schwarz inequality), that this result generalizes to produce the inequality

$$|\mathbf{y}^T A\mathbf{x}| \leqslant M(A) \qquad (9)$$

for any pair of vectors $\mathbf{y}$ and $\mathbf{x}$ for which $\mathbf{y}^T\mathbf{y} = \mathbf{x}^T\mathbf{x} = 1$, and we shall use this result in sub-section 30.2.2.

In *Unit M100 28, Linear Algebra IV*, we "measured" matrices by means of column sums rather than row sums. The measure defined here, using row sums, is a little more convenient for eigenvalue work. For symmetric matrices the two measures are, of course, always equal.

(vii) *Scaling*

There are some theoretical and computational advantages in arranging that the entries in the matrix A are of "reasonable" size, neither "too large" nor "too small". This can always be achieved by dividing every entry of A by the entry of largest magnitude, so that the magnitude of the largest entry in the new matrix is unity. The new matrix has eigenvalues which are just those of A "scaled" by the same factor. We have asked you to verify a special case of this in Exercise 4 below.

Exercises

1. Find the eigenvalues and eigenvectors, normalized so that $\mathbf{x}_r^T\mathbf{x}_r = 1$, of the matrix

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$$

2. Write down a modal matrix of Exercise 1, and verify that it is orthogonal.

3. Verify, for Exercise 1, that $X^{-1}AX = D$ in the notation of this sub-section.

4. If A is the matrix of Exercise 1 write down the scaled matrix for A, as described in the text, and calculate the eigenvalues for this scaled matrix. Verify that these are the scaled eigenvalues for A.

Solutions

1. Using techniques from *Unit 5* we know that the eigenvalues of A are the roots of

$$\det(A - \lambda I) = \lambda^2 - 4\lambda - 5 = 0,$$

that is

$$\lambda_1 = -1,\ \lambda_2 = 5.$$

We then solve the systems of equations

$$(A - \lambda_r I)\begin{bmatrix} x_{1r} \\ x_{2r} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad r = 1, 2$$

to find corresponding eigenvectors.

For $\lambda_1 = -1$ any eigenvector has the form

$$\mathbf{x}_1 = \alpha\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

where we choose α so that $\mathbf{x}_1^T\mathbf{x}_1 = 1$, that is such that $\alpha^2 + \alpha^2 = 1$ or $\alpha = \pm\dfrac{1}{\sqrt{2}}$.

Then

$$\mathbf{x}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{or} \quad \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Repeating the calculation for $\lambda_2 = 5$, we find

$$\mathbf{x}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{or} \quad \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

2. Taking the first of the two alternatives for $\mathbf{x}_1$ and $\mathbf{x}_2$ given in Solution 1, we obtain for the modal matrix

$$X = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

(There are three other possibilities arising from other choices of $\mathbf{x}_1$ and $\mathbf{x}_2$.) We easily check by matrix multiplication that

$$X^T X = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

3. Since X is orthogonal from Exercise 2, we have $X^{-1} = X^T$, so that $X^{-1}AX = X^TAX$, and we verify

$$X^TAX = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix} = D.$$

4. The scaled matrix is obtained from A by dividing each entry by 3, and is therefore

$$\begin{bmatrix} \frac{2}{3} & 1 \\ 1 & \frac{2}{3} \end{bmatrix}.$$

The eigenvalues of this matrix are the solutions of

$$\lambda^2 - \tfrac{4}{3}\lambda - \tfrac{5}{9} = 0$$

which are $\frac{5}{3}$ and $-\frac{1}{3}$. These are the eigenvalues of A divided by the same factor 3 that we used to scale A itself.

30.1.2 The Characteristic Equation and its Instabilities

In *Unit 5* we saw that the eigenvalues of a matrix A are solutions of the characteristic equation

$$\det(A - \lambda I) = 0.$$

Having found a solution λ_r of this characteristic equation, we can then solve

$$(A - \lambda_r I)\mathbf{x}_r = \mathbf{0}$$

to calculate a corresponding eigenvector. This is the method used in Exercise 1 of the previous sub-section. For any non-trivial problem, however, the method is quite uneconomic and quite unstable.

Let us first consider briefly the economy of various methods of finding the characteristic equation. Direct expansion of the characteristic polynomial $\det(A - \lambda I)$ is extremely laborious, as you will easily discover if you try this for a matrix of order as small as 6. We might, alternatively, evaluate $\det(A - \lambda I)$ for n different values of λ and solve for the n coefficients $a_0, \ldots, a_{n-1}$ the n linear simultaneous equations

$$\det(A - \lambda_i I) = a_n \lambda_i^n + a_{n-1}\lambda_i^{n-1} + \cdots + a_1\lambda_i + a_0, \qquad i = 1, 2, \ldots, n,$$

using the fact that $a_n = (-1)^n$. One can show that this requires at least $\frac{1}{3}n^4$ numerical operations. Or we may use the fact that a matrix satisfies its own characteristic equation (Cayley-Hamilton Theorem, *Unit 5*), which here means that

$$(-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I = 0 \text{ (the zero matrix).}$$

We can compute the first n powers of A and equate to zero n entries, say the entries in the first row, of the matrix on the left. This method also gives a set of linear equations for the coefficients $a_0, \ldots, a_{n-1}$. A matrix product involves n^3 operations, so that the formation of the powers of A has something like n^4 such operations, again a formidable amount of work. Next consider the accuracy of this computation. If all the elements of A are four-decimal numbers like 0.2679, we are bound to make some computer-type numerical errors in all the methods we have discussed. In our last two methods, moreover, the linear equations turn out to be very ill-conditioned, so that even if the order of A is as small as 6, small arithmetic errors produce large errors in the solution. In fact there is no known method for computing accurately the coefficients of the characteristic equation without the time-consuming expedient of using multi-length computer arithmetic to reduce the arithmetic errors in the computation.

Consider next the calculation of the solution of the characteristic equation. This is also somewhat laborious, involving iterative methods such as the Newton-Raphson method (*Unit M100 14, Sequences and Limits II*). More important, however, is the fact that the calculation of the solution of polynomial equations is often a very ill-conditioned problem, so that even small errors in the coefficients produce large errors in the roots. Consider, for example, the polynomial equation of degree 20 whose solutions are the integers 1, 2, ..., 20 which is

$$\lambda^{20} - 210\lambda^{19} + \cdots + 20! = 0.$$

If the coefficient of λ^{19} is changed by an amount 2^{-23}, less than 10^{-7}, the original solutions 16 and 17 change to the complex pair

$$\lambda_{16}, \lambda_{17} = 16.78\cdots \pm i\,2.81\ldots,$$

a substantial and most alarming change!

The practical use of the characteristic equation therefore exhibits a very large amount of induced instability. It is not inherent instability because, as we shall see in the next section, the symmetric eigenvalue problem is very well-conditioned, small changes in the matrix coefficients giving rise to only small changes in the eigenvalues.

30.1.3 Summary of Section 30.1

In this section and the Introduction we defined the terms

eigensystem	(page C5)	* *
modal matrix	(page C7)	*
scaling of matrices	(page C9)	* *

We introduced the notation

$\lambda_{\max}$	(page C8)
$M(\mathbf{x})$	(page C8)
$M(A)$	(page C8)

Main Results of Section 30.1

(i) The set V of all $n \times 1$ matrices with real entries forms a Euclidean space with inner product

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y}.$$

Every symmetric $n \times n$ matrix A has a set of eigenvectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ forming an orthonormal basis for V.

The modal matrix whose columns are $\mathbf{x}_1, \ldots, \mathbf{x}_n$ is orthogonal.

(ii) The quadratic form $\mathbf{y}^T A\mathbf{y}$ is positive definite if and only if all the eigenvalues of A are positive. It satisfies

$$|\mathbf{y}^T A\mathbf{y}| \leqslant M(A)$$

if $\mathbf{y}^T\mathbf{y} = 1$, where $m(A)$ is the largest row sum of the magnitudes of the entries of A. In addition, if $\mathbf{x}^T\mathbf{x} = 1$

$$|\mathbf{y}^T A\mathbf{x}| \leqslant M(A).$$

(iii) Multiplying all the entries in a matrix by any number (scaling) multiplies all its eigenvalues by that number.

(iv) Unless the matrix A is of order 2×2 or perhaps 3×3, the calculation of the eigenvalues of A by solving the characteristic equation is both unstable and uneconomic.

Technique

Scale a matrix so that the magnitude of the largest entry of the new matrix is 1. * *

30.2 INHERENT INSTABILITIES

30.2.0 Introduction

Following our usual plan in the numerical analysis units, we preface our discussion of numerical methods by a discussion of inherent instabilities, so that we can recognise situations in which accurate solutions are difficult or impossible to obtain however good and stable our method might be. We therefore examine the effect on the eigensolutions of small errors or uncertainties (perturbations) in the entries of A, by looking at the eigensolutions of the matrix $A + \varepsilon B$, where ε is a small number and B a symmetric matrix. Following the discussion of "scaling" in sub-section 30.1.1 we can assume, without loss of generality, that all the entries in A and B have magnitude less than or equal to 1.

It is reasonable to suppose that for sufficiently small ε the eigenvalue $\lambda_r(\varepsilon)$ of $A + \varepsilon B$ corresponding to the eigenvalue λ_r of A will not differ very much from λ_r. In fact, it can be shown that $\lambda_r(\varepsilon)$ can be written as a Taylor series in powers of ε.

$$\lambda_r(\varepsilon) = \lambda_r + k_r\varepsilon + \text{terms in } \varepsilon^2, \varepsilon^3, \ldots, \tag{1}$$

where the number k_r depends on A and B but not on ε. We call $k_r\varepsilon$ the "first-order" perturbation in the eigenvalue λ_r, and we hope to find a computable expression for k_r. We shall not try to estimate the "higher-order" perturbation terms in (1).

It can also be shown that every entry in $\mathbf{x}_r(\varepsilon)$, the eigenvector corresponding to $\lambda_r(\varepsilon)$, can be written as a Taylor series in ε, having the form

$$\mathbf{x}_r(\varepsilon) = \mathbf{x}_r + \varepsilon\mathbf{z}_r + \text{terms in } \varepsilon^2, \varepsilon^3 \ldots; \tag{2}$$

where $\mathbf{x}_r$ is the normalized eigenvector of A corresponding to λ_r, and $\mathbf{z}_r$ is a vector depending on A and B but not on ε. To fix the length (and sense) of the eigenvector we impose the condition

$$\mathbf{x}_r \cdot \mathbf{x}_r(\varepsilon) = 1$$

so that (2) gives

$$\mathbf{x}_r \cdot \mathbf{x}_r + \varepsilon\mathbf{x}_r \cdot \mathbf{z}_r + \cdots = 1 \tag{3}$$

Since this holds for all ε, it implies that the first-order perturbation $\varepsilon\mathbf{z}_r$ (and all higher-order perturbations) in $\mathbf{x}_r$ is orthogonal to $\mathbf{x}_r$.

From the definition of an eigenvalue problem, we have

$$(A + \varepsilon B)(\mathbf{x}_r + \varepsilon\mathbf{z}_r + \cdots) = (\lambda_r + \varepsilon k_r + \cdots)(\mathbf{x}_r + \varepsilon\mathbf{z}_r + \cdots)$$

Multiplying out the products and writing down only the terms of degree 1 or less in ε, we find

$$A\mathbf{x}_r + \varepsilon(B\mathbf{x}_r + A\mathbf{z}_r) + \cdots = \lambda_r\mathbf{x}_r + \varepsilon(k_r\mathbf{x}_r + \lambda_r\mathbf{z}_r) + \cdots$$

(this uses the rule for multiplying power series, Theorem I-31 on page K665). The two power series represent the same function of ε and we therefore have

$$\begin{aligned} A\mathbf{x}_r &= \lambda_r\mathbf{x}_r \\ B\mathbf{x}_r + A\mathbf{z}_r &= k\mathbf{x}_r + \lambda_r\mathbf{z}_r \end{aligned} \tag{4}$$

The first equation tells us nothing new, but from Equation (4) we can proceed to determine the first-order perturbations in the eigenvalue λ_r and in the corresponding eigenvector $\mathbf{x}_r$.

30.2.1 The Eigenvalues

To find k_r we have to eliminate the vector $\mathbf{z}_r$ in Equation (4) of the previous sub-section, and this is effected immediately by taking the inner product of both sides of this equation with the eigenvector $\mathbf{x}_r$. When we do this, the term $\mathbf{x}_r \cdot A\mathbf{z}_r$ simplifies as follows, since A is symmetric:

$$\begin{aligned} \mathbf{x}_r \cdot A\mathbf{z}_r &= (A\mathbf{x}_r) \cdot \mathbf{z}_r \\ &= \lambda_r \mathbf{x}_r \cdot \mathbf{z}_r \\ &= 0; \end{aligned}$$

and since $\mathbf{x}_r$ is orthogonal to $\mathbf{z}_r$, the last term on the right $\mathbf{x}_r \cdot \lambda_r \mathbf{z}_r$ is also 0. Our formula therefore becomes

$$\mathbf{x}_r \cdot B\mathbf{x}_r + 0 = k_r \mathbf{x}_r \cdot \mathbf{x}_r + 0$$

which simplifies to

$$k_r = \mathbf{x}_r \cdot B\mathbf{x}_r \qquad (1)$$

since $\mathbf{x}_r$ is normalized.

Using Equation (8) of sub-section 30.1.1 we can then write

$$|\varepsilon||k_r| \leqslant |\varepsilon| M(B) \leqslant |\varepsilon| n \qquad (2)$$

since the matrix B is scaled and has n entries in each row.

Now from Equation 1 of sub-section 30.2.0 we have

$$|\lambda_r(\varepsilon) - \lambda_r| \leqslant |\varepsilon||k_r| \qquad (3)$$

Hence from Equations (2) and (3) we obtain

$$\begin{aligned} |\lambda_r(\varepsilon) - \lambda_r| &\leqslant |\varepsilon| M(B) \qquad (4) \\ &= M(\varepsilon B) \qquad (5) \end{aligned}$$

Equations (2), (3) and (4) show that small changes, of magnitude at most ε, in each of the n^2 entries of A produce "first-order" changes in the eigenvalue that are at most only n times as large: *the determination of the eigenvalues of a symmetric matrix is quite a well-conditioned problem.*

Example

Consider the matrix $A = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$

which we can show has the normalized eigensolutions

$$\lambda_1 = 1.5, \quad \mathbf{x}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\lambda_2 = 0.5, \quad \mathbf{x}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and consider the perturbation $\varepsilon B = \varepsilon \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

Our preceding analysis tells us that under such a perturbation the greatest possible "first-order" change in any eigenvalue is $n\varepsilon$ with $n = 2$; in fact we have here

$$\varepsilon \mathbf{x}_1{}^T B \mathbf{x}_1 = \varepsilon \frac{1}{\sqrt{2}} [1 \quad 1] \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2\varepsilon.$$

This is here the *exact* change; for the characteristic equation of $A + \varepsilon B$ is

$$(1 + \varepsilon - \lambda)^2 - (0.5 + \varepsilon)^2 = 0$$

or $\quad (1 + \varepsilon - \lambda) = \pm (0.5 + \varepsilon)$

giving $\quad \lambda_1, \lambda_2 = (1 + \varepsilon) \pm (0.5 + \varepsilon),$

so that $\quad \lambda_1 = 1.5 + 2\varepsilon, \ \lambda_2 = 0.5.$

Exercise

With the A of the previous example, find a symmetric matrix B, all of whose entries are either $+1$ or -1, for which the eigenvalue of $A + \varepsilon B$, corresponding to λ_2 of A, differs from λ_2 by an amount of modulus approximately 2ε. Verify that this is the exact change.

Solution

The eigenvector corresponding to λ_2 is

$$\mathbf{x}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Then

$$\varepsilon \mathbf{x}_2^T B \mathbf{x}_2 = \tfrac{1}{2}\varepsilon[1 \quad -1]\begin{bmatrix} \pm 1 & 1 \\ 1 & \pm 1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

where, obviously without loss of generality, we have taken the off-diagonal terms of B to be $+1$. If we call the diagonal terms α and β, we find

$$\varepsilon \mathbf{x}_2^T B \mathbf{x}_2 = \tfrac{1}{2}\varepsilon(\alpha + \beta - 2),$$

so that for *maximum* size we take $\alpha = \beta = -1$, and the approximate change in the eigenvalue has magnitude 2ε.

The characteristic equation of $A + \varepsilon B$ is then

$$(1 - \varepsilon - \lambda)^2 - (0.5 + \varepsilon)^2 = 0$$

or

$$(1 - \varepsilon - \lambda) = \pm(0.5 + \varepsilon)$$

so that

$$\lambda_1, \lambda_2 = 1 - \varepsilon \pm (0.5 + \varepsilon).$$

Then

$$\lambda_2 = 0.5 - 2\varepsilon,$$

and our *approximate* change is the *exact* change.

30.2.2 The Eigenvectors

To estimate the change in an eigenvector, say $\mathbf{x}_r$, produced by the perturbation εB in the matrix A, we go back to Equation (4) of sub-section 30.2.0, which is

$$B\mathbf{x}_r + A\mathbf{z}_r = k_r \mathbf{x}_r + \lambda_r \mathbf{z}_r.$$

This time we want to eliminate k_r; so we take the scalar product with any eigenvector $\mathbf{x}_s$ of A *other* than $\mathbf{x}_r$. This gives

$$\mathbf{x}_s \cdot B\mathbf{x}_r + \mathbf{x}_s \cdot A\mathbf{z}_r = \lambda_r \mathbf{x}_s \cdot \mathbf{z}_r \qquad (1)$$

Since A is symmetric, the second term on the left is equal to

$$(A\mathbf{x}_s) \cdot \mathbf{z}_r = \lambda_s \mathbf{x}_s \cdot \mathbf{z}_r,$$

and so Equation (1) can be simplified to give

$$\mathbf{x}_s \cdot \mathbf{z}_r = \frac{\mathbf{x}_s \cdot B\mathbf{x}_r}{\lambda_r - \lambda_s}. \qquad (2)$$

This is true for every $s \neq r$. The first-order perturbation in the eigenvector $\mathbf{x}_r$ produced by a small perturbation εB in the matrix A is therefore given (see Equation 7-46 on page K278) by

$$\begin{aligned}\varepsilon \mathbf{z}_r &= \varepsilon \sum_{s=1}^{n} (\mathbf{x}_s \cdot \mathbf{z}_r)\mathbf{x}_s \\ &= \varepsilon \sum_{\substack{s=1 \\ (s \neq r)}}^{n} \left(\frac{\mathbf{x}_s \cdot B\mathbf{x}_r}{\lambda_r - \lambda_s}\right)\mathbf{x}_s \end{aligned} \qquad (3)$$

since $\mathbf{x}_s \cdot \mathbf{z}_r$ is 0 if $s = r$ and is given by Equation (2) if $s \neq r$.

By Equation (9) of sub-section 30.1.1, the number $\mathbf{x}_s \cdot B\mathbf{x}_r$ has magnitude not exceeding $M(B)$, so that the coefficient of $\mathbf{x}_s$ in (3) has magnitude not exceeding

$$\frac{|\varepsilon|\, M(B)}{|\lambda_r - \lambda_s|} \leqslant \frac{n\,|\varepsilon|}{|\lambda_r - \lambda_s|} \qquad (4)$$

since B is scaled.

We see that the coefficient of $\mathbf{x}_s$ is not large provided that $\lambda_r - \lambda_s$ is not small. On the other hand, if one or more of the other eigenvalues of A is close to λ_r then the perturbation in the eigenvector $\mathbf{x}_r$ is not necessarily small in virtue of the small perturbations εB in A. In other words, *eigenvectors corresponding to eigenvalues λ_r and λ_s which are nearly equal may be ill-conditioned,* and the degree of ill-conditioning is proportional to $(\lambda_r - \lambda_s)^{-1}$. This result is perhaps not surprising when we look at the limiting case when λ_r and λ_s become exactly equal. For then there are two orthogonal eigenvectors $\mathbf{x}_r$ and $\mathbf{x}_s$ with the same eigenvalue $\lambda_r = \lambda_s$, and every non-zero vector in the eigenspace $\langle \mathbf{x}_r, \mathbf{x}_s \rangle$ is also an eigenvector. In other words, when $\lambda_r = \lambda_s$ we cannot determine uniquely the directions of the corresponding eigenvectors, and it is not therefore surprising that we have difficulty in determining these directions when λ_r is close to λ_s.

Example

The matrix

$$A = \begin{bmatrix} 1 & 0.01 \\ 0.01 & 1 \end{bmatrix}$$

has the normalized eigensolutions

$$\lambda_1 = 1.01, \mathbf{x}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \simeq \begin{bmatrix} 0.71 \\ 0.71 \end{bmatrix},$$

$$\lambda_2 = 0.99, \mathbf{x}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \simeq \begin{bmatrix} 0.71 \\ -0.71 \end{bmatrix}.$$

We consider the effect of a perturbation εB, with $\varepsilon = 0.04$ and

$$B = \begin{bmatrix} 1 & 0.5 \\ 0.5 & -1 \end{bmatrix},$$

and to this end we calculate the eigensolutions of the matrix

$$A + \varepsilon B = \begin{bmatrix} 1.04 & 0.03 \\ 0.03 & 0.96 \end{bmatrix}$$

The eigenvalues are 1.05 and 0.95, not very different from those of A in accordance with our expectations.

The normalized eigenvectors are multiples of the vectors

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ -3 \end{bmatrix},$$

and they are clearly very different from those of A. In fact if we take the first eigenvector to be $k\begin{bmatrix}3\\1\end{bmatrix}$, and choose k, as suggested in sub-section 30.2.0, so that we can write

$$k\begin{bmatrix}3\\1\end{bmatrix} = \mathbf{x}_1 + \alpha\mathbf{x}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} + \frac{\alpha}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$$

with the coefficient of $\mathbf{x}_1$ equal to 1, we find

$$k = \frac{1}{\frac{1}{\sqrt{2}}[1 \quad 1]\begin{bmatrix}3\\1\end{bmatrix}} = \frac{1}{2\sqrt{2}}$$

and

$$\alpha = k\frac{1}{\sqrt{2}}[1 \quad -1]\begin{bmatrix}3\\1\end{bmatrix} = \frac{1}{2}.$$

This means that the perturbed eigenvector corresponding to $\mathbf{x}_1$ has acquired a large component, of magnitude $\frac{1}{2}$, in the direction of $\mathbf{x}_2$!

You should realize that the formulas (1) of sub-section 30.2.1 and (3) of this sub-section do not here give very accurate expressions for the perturbations. This is because the number $\lambda_1 - \lambda_2 = 0.02$ is small even compared with our particular ε, so that "higher-order" effects are here important. The general conclusions, of course, are not affected. Even "first-order" perturbations are large in eigenvectors corresponding to close eigenvalues, and this is enough to guarantee the ill-conditioning of these eigenvectors. Higher-order effects, on the other hand, cannot affect our conclusion about the good conditioning of the eigenvalues. In fact we can show, by more advanced techniques which we shall not give here, that Equation (2) of sub-section 30.2.1, giving a bound for the first-order perturbation of an eigenvalue, is true without the qualification "first-order". In other words, for any perturbation B, however large, no eigenvalue of $A + \varepsilon B$ can differ from some eigenvalue of A by more than $M(B)$ in magnitude!

Exercises

1. In the example of this sub-section, show that the perturbation in the eigenvalue λ_1, obtained from Formula (1) of sub-section 30.2.1 is smaller than the true perturbation, but that the latter still satisfies Equation (2) of sub-section 30.2.1 and verifies the last remark of the text.

2. If the perturbation of the example of this sub-section were applied to the matrix of Exercise 1 of sub-section 30.1.1, why is the effect very small on both eigensolutions?

Solutions

1. The first-order perturbation in λ_1 is

$$\varepsilon\mathbf{x}_1^T B\mathbf{x}_1 = 0.04\frac{1}{\sqrt{2}}[1 \quad 1]\begin{bmatrix}1 & 0.5\\0.5 & -1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} = 0.02,$$

whereas the exact change in λ_1 is 0.04. This exact change is smaller in magnitude than $\varepsilon M(B)$, which is $0.04(1.5) = 0.06$.

2. The eigenvalues of the given matrix are $\lambda_1 = -1$, $\lambda_2 = 5$, which are well separated. Hence the eigenvectors, as well as the eigenvalues, are well-conditioned, affected only slightly by a small perturbation in the matrix.

30.2.3 Summary of Section 30.2

In this section we defined the terms

perturbation of a matrix	(page C13)
first-order perturbation of an eigenvalue	(page C13)
first-order perturbation of an eigenvector	(page C13)

We introduced the notation

$\lambda_r(\varepsilon)$	(page C13)
$\mathbf{x}_r(\varepsilon)$	(page C13)

Main Results of Section 30.2

(i) The first-order perturbation εk_r of the eigenvalue λ_r of a symmetric $n \times n$ matrix A caused by the small perturbation εB of A is bounded as follows:

$$|\varepsilon||k_r| \leqslant |\varepsilon| M(B) \leqslant |\varepsilon| n$$

In other words, the determination of the eigenvalues of a symmetric matrix is quite a well-conditioned problem.

(ii) The first-order perturbation $\varepsilon \mathbf{z}_r$ of the eigenvector $\mathbf{x}_r$ of A satisfies

$$\varepsilon \mathbf{z}_r = \varepsilon \sum_{\substack{s=1 \\ (s \neq r)}}^{n} \left(\frac{\mathbf{x}_s \cdot B\mathbf{x}_r}{\lambda_r - \lambda_s} \right) \mathbf{x}_s$$

and

$$\frac{|\varepsilon||\mathbf{x}_s \cdot B\mathbf{x}_r|}{|\lambda_r - \lambda_s|} \leqslant \frac{|\varepsilon| M(B)}{|\lambda_r - \lambda_s|} \leqslant \frac{n|\varepsilon|}{|\lambda_r - \lambda_s|}$$

In other words, eigenvectors corresponding to eigenvalues λ_r and λ_s which are nearly equal may be ill-conditioned.

30.3 ITERATIVE METHODS FOR INDIVIDUAL EIGENSOLUTIONS

30.3.0 Introduction

This section and Section 30.4 present useful practical methods for computing eigensolutions. In this section we examine iterative methods, which are most convenient if we want just one, or perhaps a very few, of the eigensolutions. We shall discuss two important iterative methods. The first is called *direct iteration* and this, as we shall see, can produce only two eigensolutions, those for which the corresponding eigenvalues are farthest from zero in the positive or negative directions. The second is called *inverse iteration*, and this can produce any eigensolution and is therefore, in some senses, a more powerful method.

30.3.1 Direct Iteration: Theory

We have already observed (in sub-section 30.1.1) that any vector $\mathbf{y}$ in V can be expressed as a linear combination of eigenvectors in the form

$$\mathbf{y} = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_n \mathbf{x}_n \tag{1}$$

What we try to do with iterative methods is to operate on $\mathbf{y}$ in such a way that one of the coefficients α_r is magnified at the expense of all the others, so that we transform $\mathbf{y}$ into a good approximation to some eigenvector.

The simplest operation uses the linear transformation A itself. We considered the effect of applying such a transformation repeatedly in *Unit 5*, sub-section 5.3.2, but a little generalization produces more useful techniques. We therefore consider a linear transformation with the matrix $A - pI$, where p is some real number, and observe that

$$(A - pI)\mathbf{x}_i = A\mathbf{x}_i - p\mathbf{x}_i = (\lambda_i - p)\mathbf{x}_i, \tag{2}$$

where $\mathbf{x}_i$ and λ_i represent an eigensolution of A. It follows that the corresponding eigensolution of $A - pI$ is

$$\mathbf{x}_i \text{ and } \lambda_i - p.$$

From Equations (1) and (2) we find

$$(A - pI)\mathbf{y} = \alpha_1(\lambda_1 - p)\mathbf{x}_1 + \alpha_2(\lambda_2 - p)\mathbf{x}_2 + \cdots + \alpha_n(\lambda_n - p)\mathbf{x}_n,$$
$$(A - pI)^2\mathbf{y} = \alpha_1(\lambda_1 - p)^2\mathbf{x}_1 + \alpha_2(\lambda_2 - p)^2\mathbf{x}_2 + \cdots + \alpha_n(\lambda_n - p)^2\mathbf{x}_n,$$

and in general

$$(A - pI)^r\mathbf{y} = \alpha_1(\lambda_1 - p)^r\mathbf{x}_1 + \alpha_2(\lambda_2 - p)^r\mathbf{x}_2 + \cdots + \alpha_n(\lambda_n - p)^r\mathbf{x}_n \tag{3}$$

Now as r increases the right-hand side of (3) is increasingly "dominated" by the term $\alpha_k(\lambda_k - p)^r\mathbf{x}_k$, for which $|\lambda_k - p|$ is the largest of all the $|\lambda_i - p|$, $(i = 1, \ldots, n)$ provided only that $\alpha_k \neq 0$ and that the k so defined is unique.

We can then write (3) in the form

$$(A - pI)^r\mathbf{y} = (\lambda_k - p)^r \left\{\alpha_1 \left(\frac{\lambda_1 - p}{\lambda_k - p}\right)^r \mathbf{x}_1 + \cdots + \alpha_k \mathbf{x}_k + \cdots + \alpha_n \left(\frac{\lambda_n - p}{\lambda_k - p}\right)^r \mathbf{x}_n\right\} \tag{4}$$

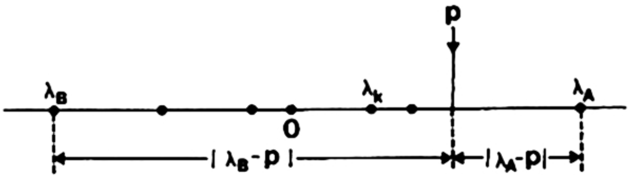
and as r increases every term in the bracket other than $\alpha_k \mathbf{x}_k$ gets smaller and smaller, and $(A - pI)^r\mathbf{y}$ becomes more and more nearly proportional to the eigenvector $\mathbf{x}_k$. (We saw an example of this in the last sub-section of *Unit 5*.)

The same calculation also gives us the eigenvalue λ_k. For sufficiently large r we have

$$(A - pI)^{r+1}\mathbf{y} = (\lambda_k - p)^{r+1}\{\alpha_k \mathbf{x}_k + \text{small vectors}\}$$
$$(A - pI)^{r}\mathbf{y} = (\lambda_k - p)^{r}\{\alpha_k \mathbf{x}_k + \text{small vectors}\},$$

and it follows that as r gets larger and larger the ratio of any corresponding non-zero entries in the vectors $(A - pI)^{r+1}\mathbf{y}$ and $(A - pI)^{r}\mathbf{y}$ gets closer and closer to $\lambda_k - p$. With a prescribed p we therefore compute a close approximation to the eigenvalue λ_k.

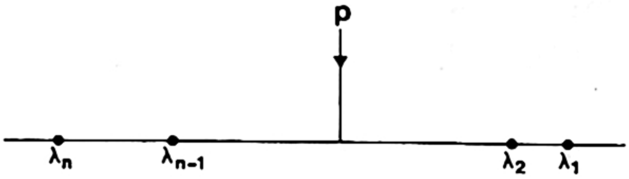
This method of computing an eigensolution is called *direct iteration* with $A - pI$. Two important questions are: how do we choose the number p, and what particular eigenvalues λ_k can we actually compute by this method? To answer these questions consider the eigenvalues located at points on the real line, shown below.



Then the values of $|\lambda_k - p|$ are the distances of these points from the point p on the real line. If p is closer to the largest eigenvalue λ_A than to the smallest λ_B, then the largest value of $|\lambda_k - p|$ is $|\lambda_B - p|$. If p is closer to λ_B, then the largest $|\lambda_k - p|$ is $|\lambda_A - p|$. It is clear that the only eigensolutions to which direct iteration will converge are those with eigenvalues λ_A and λ_B, that is to the eigensolutions corresponding to the two *extreme* eigenvalues.

This answers both our questions, though of course a decision about what p to choose in order to guarantee convergence to λ_A say, rather than to λ_B, will need either some prior knowledge of the approximate values of λ_A or λ_B, or a process of trial and error.

But there is another interesting thing about the choice of p which concerns the *rate* of convergence of our iterative process. To examine this let us change our notation, labelling the eigenvalues so that $\lambda_1 > \lambda_2 > \cdots > \lambda_n$, as shown below.



Let us suppose that we want to calculate λ_n, the smallest eigenvalue. The relevant equation, from (4), is

$$(A - pI)^{r}\mathbf{y} = (\lambda_n - p)^{r}\left\{\alpha_n \mathbf{x}_n + \alpha_1\left(\frac{\lambda_1 - p}{\lambda_n - p}\right)^{r}\mathbf{x}_1 + \cdots + \alpha_{n-1}\left(\frac{\lambda_{n-1} - p}{\lambda_n - p}\right)^{r}\mathbf{x}_{n-1}\right\}.$$

For convergence, as we have seen, we must choose p so that

$$|\lambda_1 - p| < |\lambda_n - p|;$$

this ensures that $|\lambda_n - p|$ is the largest of the $|\lambda_i - p|$. Provided $\alpha_n \neq 0$, the *rate* of convergence clearly depends on the rates at which the various $|\lambda_i - p|/|\lambda_n - p|$ get smaller as r increases, and we would like to choose p so that the *largest* of these numbers is as *small* as possible.

Now the " worst offenders " are clearly λ_1 and λ_{n-1}, and the best choice for p is that for which these give equal and opposite ratios, that is

$$\frac{\lambda_1 - p}{\lambda_n - p} = -\frac{\lambda_{n-1} - p}{\lambda_n - p}, \quad \text{or} \quad \lambda_1 - p = -(\lambda_{n-1} - p),$$

giving

$$p = \tfrac{1}{2}(\lambda_1 + \lambda_{n-1}).$$

Any other choice would make one of these ratios larger than it need be!

By precisely the same argument we can show that the best choice of p for convergence to λ_1, the largest eigenvalue, is

$$p = \tfrac{1}{2}(\lambda_n + \lambda_2).$$

Example

Suppose that we have a matrix of order 3, with eigenvalues 5, 3 and -1. Taking $p = 0$, that is iterating directly with A, we converge to the eigensolution with $\lambda = 5$, and the ratios which govern the rate of convergence are $\tfrac{3}{5}$ and $\tfrac{1}{5}$. The choice

$$p = \tfrac{1}{2}(3 - 1) = 1$$

gives

$$\lambda_1 - p = 4, \lambda_2 - p = 2, \lambda_3 - p = -2,$$

and we still converge to the largest eigenvalue but at a rate governed by the ratios

$$\left|\frac{\lambda_2 - p}{\lambda_1 - p}\right|, \left|\frac{\lambda_3 - p}{\lambda_1 - p}\right|,$$

both of which are $\tfrac{1}{2}$. Since $(\tfrac{1}{2})^r$ decreases considerably faster than $(\tfrac{3}{5})^r$, as r increases, we have achieved a much better rate of convergence.

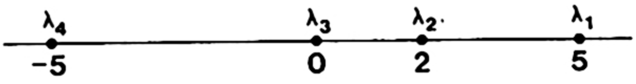
It is only fair to remark here that a good choice of p requires even more knowledge about the distribution of the eigenvalues. But trial and error can still play its part, and in early work with digital computers this was quite a powerful method, since p could be varied at will and the effects of convergence could be observed on something like a television screen.

Exercises

1. A certain matrix has eigenvalues -5, 0, 2 and 5. For what values of p would we, by direct iteration with $A - pI$, obtain fastest convergence (i) to the largest eigenvalue, (ii) to the smallest? What, in each case, are the ratios governing the rate of convergence?
2. Can you see how you might get both the largest and smallest eigenvalues, and the corresponding eigenvectors, almost simultaneously by direct iteration with A(that is, $p = 0$), in this particular case when $\lambda_1 = -\lambda_n$?

Solutions

1. Let us represent the eigenvalues diagrammatically.



(i) For convergence to the largest eigenvalue, which is $\lambda_1 = 5$, we choose p to be the average of the largest and smallest unwanted eigenvalues, which is

$$\tfrac{1}{2}(-5 + 2) = -\tfrac{3}{2}.$$

The ratios $\frac{\lambda_k - p}{\lambda_1 - p}$ governing the rate of convergence are $\frac{-7}{13}, \frac{3}{13}, \frac{7}{13}$, representing a reasonable rate of convergence,

(ii) For convergence to the smallest eigenvalue, which is $\lambda_4 = -5$, we choose p to be the average of 0 and 5, which is $\frac{5}{2}$. The relevant ratios are $\frac{5}{15}, \frac{1}{15}, \frac{-5}{15}$, and we get faster convergence than in the previous case.

2. Direct iteration with A will give, for sufficiently large r,

$$A^r\mathbf{y} = \lambda_1^r\{(\alpha_1\mathbf{x}_1 + (-1)^r\alpha_n\mathbf{x}_n) + \text{negligible vectors}\}$$

So λ_1^2, giving λ_1 and $-\lambda_1 = \lambda_n$, is the ratio of corresponding entries of $A^{r+2}\mathbf{y}$ and $A^r\mathbf{y}$. Having obtained λ_1, we can then write

$$\frac{A^r\mathbf{y}}{\lambda_1^r} = \alpha_1\mathbf{x}_1 \pm \alpha_n\mathbf{x}_n$$

$$\frac{A^{r+1}\mathbf{y}}{\lambda_1^{r+1}} = \alpha_1\mathbf{x}_1 \mp \alpha_n\mathbf{x}_n$$

from which we easily find multiples of the required vectors $\mathbf{x}_1$ and $\mathbf{x}_n$.

30.3.2 Direct Iteration: Practice

It is quite uneconomic to compute successive powers of the matrix $A - pI$, a procedure which appeared to be needed in our discussion of the theory of direct iteration. We prefer to express the method in terms of an iterative algorithm of the form

$$\mathbf{y}^{(r+1)} = (A - pI)\mathbf{y}^{(r)}, \ \mathbf{y}^{(0)} \text{ arbitrary},$$

which you can easily see is identical, at least in theory, with the statement

$$\mathbf{y}^{(r+1)} = (A - pI)^{r+1}\mathbf{y}^{(0)}$$

implied in the previous sub-section.

In practice we go just a little further, in the attempt, already mentioned in sub-section 30.1.1 (vii), to keep all our numbers of reasonable size. We do this by dividing each successive vector by its entry of largest magnitude, so that the new vector has the magnitude of its largest entry equal to unity. This produces the algorithm for direct iteration in the form

$$\begin{aligned}&\mathbf{y}^{(0)} \text{ arbitrary (with largest entry} = 1)\\ &\mathbf{z}^{(r+1)} = (A - pI)\mathbf{y}^{(r)}\\ &\mathbf{y}^{(r+1)} = \mathbf{z}^{(r+1)}/m(\mathbf{z}^{(r+1)}),\end{aligned} \qquad (1)$$

where $m(\mathbf{z})$ means the entry in $\mathbf{z}$ with largest magnitude. (Note that $|m(\mathbf{z})| = M(\mathbf{z})$.)

Apart from keeping the numbers to reasonable size this device has two other advantages. First, $m(\mathbf{z}^{(r+1)})$ is the best current estimate of the number $\lambda_1 - p$ or $\lambda_n - p$, to which we are converging, at least for an r large enough to ensure that successive vectors have the largest numbers in the *same* entry. Second, and more important, this process answers another question: what happens if the initial guess $\mathbf{y}^{(0)}$ is orthogonal to the eigenvector we are hoping to find, that is, if α_1 (or α_n) = 0 in the equation

$$\mathbf{y}^{(0)} = \sum_{r=1}^{n}\alpha_r\mathbf{x}_r\ ?$$

In theory, and certainly with *exact* arithmetic, we could converge to some other solution! In practice we want to be sure that we have converged to one of the extreme solutions, and our scaling process virtually ensures this.

For with computer arithmetic the division $\mathbf{z}^{(r+1)}/m(\mathbf{z}^{(r+1)})$ will sooner or later (and usually sooner!) involve rounding errors, so that the *computed* $\bar{\mathbf{y}}^{(r+1)}$ will equal the true $\mathbf{y}^{(r+1)}$ plus a vector of rounding errors. This will include a small multiple of the required $\mathbf{x}_1$(or $\mathbf{x}_n$), and this multiple is steadily magnified as r increases. Convergence will be slow at first but ultimately certain, and here rounding errors act to our advantage!

Example

For the matrix $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 3 & 1 & 4 \end{bmatrix}$

we start with

$$\mathbf{y}^{(0)} = [1 \quad 1 \quad 1]^T,$$

and (taking $p = 0$) produce the following results in three-digit floating point arithmetic. The column labelled λ_1 is $m(\mathbf{z}^{(r)})$ and is the current estimate of the dominant eigenvalue. (You need check only the first one or two steps.)

r	$\mathbf{z}^{(r)T}$	$\mathbf{y}^{(r)T}$	λ_1
0	[1 1 1]	[1 1 1]	
1	[6 1 8]	[0.750 0.125 1]	8
2	[4.63 1.63 6.38]	[0.726 0.255 1]	6.38
3	[4.71 1.47 6.43]	[0.733 0.229 1]	6.43
4	[4.70 1.50 6.43]	[0.731 0.233 1]	6.43
5	[4.70 1.50 6.43]	[0.731 0.233 1]	6.43

To three-figure accuracy we have

$$\mathbf{z}^{(5)} = A\mathbf{y}^{(4)}$$

and

$$\mathbf{y}^{(5)} = \mathbf{z}^{(5)}/6.43 = \mathbf{y}^{(4)}.$$

To this accuracy 6.43 and $\mathbf{y}^{(5)}$ is an eigensolution, and we can confidently expect that it is the dominant eigensolution.

Exercises

1. (i) Given that the eigenvalues of A in the last example are approximately 6.4, -1.3 and -0.1, what value of p would you use for fastest convergence to the dominant eigensolution by iterating with $A - pI$.
   (ii) Taking $p = -1$ for simplicity, perform two steps of the process, using three-digit arithmetic and starting as before with $\mathbf{y}^{(0)} = [1 \quad 1 \quad 1]^T$.
   (iii) What is the best estimate of the dominant eigenvalue of A given by this calculation?

2. The matrix

$$A = \begin{bmatrix} 1.3 & 2.0 & 0.4 \\ 2.0 & 1.9 & -1.6 \\ 0.4 & -1.6 & 3.1 \end{bmatrix}$$

has eigenvalues $\lambda_1 = 4.5$, $\lambda_2 = 2.7$, $\lambda_3 = -0.9$. An eigenvector corresponding to λ_1 is

$$\mathbf{x}_1 = [0.5 \quad 1.0 \quad -1.0]^T.$$

(i) Show that the vector

$$\mathbf{y}^{(0)} = [0.4 \quad 0.8 \quad 1.0]^T$$

is orthogonal to the eigenvector $\mathbf{x}_1$, so that $\alpha_1 = 0$ in

$$\mathbf{y}^{(0)} = \sum_{r=1}^{3} \alpha_r \mathbf{x}_r .$$

(ii) Show, however, that if we use three-decimal floating-point arithmetic the iterative method of this sub-section, with $p = 0$, *will* converge to the dominant eigensolution λ_1 and $\mathbf{x}_1$.

Solutions

1. (i) The best value of p is half-way between -1.3 and -0.1,

so that $p = -0.7$.

(ii) Taking $p = -1$, we have

$$A - pI = \begin{bmatrix} 3 & 1 & 3 \\ 1 & 0 & 1 \\ 3 & 1 & 5 \end{bmatrix}$$

and two steps of the iteration give

r	$\mathbf{z}^{(r)T}$	$\mathbf{y}^{(r)T}$	$\lambda_1 + 1$
0	[1 1 1]	[1 1 1]	
1	[7 2 9]	[0.778 0.222 1]	9
2	[5.55 1.78 7.55]	[0.735 0.236 1]	7.55

(iii) At this stage the vector appears to be converging nicely, and the current best estimate for the required eigenvalue is $\lambda - p = 7.55$.

With $p = -1$ this gives $\lambda = 6.55$. We are converging, as we would expect, just a little faster than we did in the previous example.

2. (i) With $\mathbf{x}_1^T = [0.5 \quad 1.0 \quad -1.0]$, $\mathbf{y}^{(0)} = [0.4 \quad 0.8 \quad 1.0]^T$, the inner product

$$\mathbf{x}_1^T \mathbf{y}^{(0)} = (0.5)(0.4) + (1.0)(0.8) + (-1.0)(1.0) = 0$$

Hence $\mathbf{y}^{(0)}$ is orthogonal to $\mathbf{x}_1$, and $\alpha_1 = 0$ in

$$\mathbf{y}^{(0)} = \sum_{r=1}^{3} \alpha_r \mathbf{x}_r .$$

(ii) With $\mathbf{y}^{(0)}$ as first approximation the iteration proceeds as follows:

$$A = \begin{bmatrix} 1.3 & 2.0 & 0.4 \\ 2.0 & 1.9 & -1.6 \\ 0.4 & -1.6 & 3.1 \end{bmatrix}$$

r	$\mathbf{z}^{(r)T}$	$\mathbf{y}^{(r)T}$
0	[0.4 0.8 1.0]	[0.4 0.8 1.0]
1	[2.52 0.72 1.98]	[1 0.286 0.786]

There are already rounding errors in the formation of $\mathbf{y}^{(1)}$ from $\mathbf{z}^{(1)}$, so that a multiple of $\mathbf{x}_1$ is now present in $\mathbf{y}^{(1)}$, and we shall converge to $\mathbf{x}_1$ since $\lambda_1 = 4.5$ is the largest eigenvalue and we are taking $p = 0$. (On the O.U. computer we get the answer correct to three decimals after 56 iterations.)

30.3.3 Direct Iteration: Induced Instabilities

Direct iteration for the extreme eigensolutions has virtually no induced instability. It will always produce reasonably accurate eigenvalues, and the accuracy of the computed eigenvectors will depend, as one would expect, on the inherent instability of the problem. We can see this with the help of a backward error analysis, similar in spirit to the one we used in *Unit 8* for the Gauss elimination method.

The iteration method with $p = 0$ is described theoretically by the equations

$$\left.\begin{aligned} \mathbf{z}^{(r+1)} &= A\mathbf{y}^{(r)} \\ \mathbf{y}^{(r+1)} &= \mathbf{z}^{(r+1)}/m(\mathbf{z}^{(r+1)}) \end{aligned}\right\} \tag{1}$$

that is, by

$$k^{(r+1)}\mathbf{y}^{(r+1)} = A\mathbf{y}^{(r)}$$

where

$$k^{(r+1)} = m(\mathbf{z}^{(r+1)})$$

Owing to rounding errors in the computation of $A\mathbf{y}^{(r)}$, the calculation will actually give

$$k^{(r+1)}\mathbf{y}^{(r+1)} = A\mathbf{y}^{(r)} + \mathbf{f}^{(r)} \tag{2}$$

where $\mathbf{f}^{(r)}$ is a vector of rounding errors.

Now we would like to be able to write this equation in the form

$$k^{(r+1)}\mathbf{y}^{(r+1)} = (A + Z^{(r)})\mathbf{y}^{(r)} \tag{3}$$

so that we could say that the vectors we have actually produced at this stage were obtained from the perturbed matrix $A + Z^{(r)}$ rather than from A itself.

We would then like some estimate of this perturbing matrix $Z^{(r)}$. Now the vector $\mathbf{y}^{(r)}$ has been normalized so that its largest entry is 1. Suppose that this largest entry is the sth, and let

$$\mathbf{e}_s = [0 \quad 0 \cdots 0 \quad 1 \quad 0 \cdots 0]^T$$

be the vector with zeros everywhere except for a 1 in the sth position, so that $\mathbf{e}_s$ is in fact the sth column of the unit matrix I. Now write

$$Z^{(r)} = \mathbf{f}^{(r)}\mathbf{e}_s^T = \begin{bmatrix} 0 \cdots 0 & f_1^{(r)} & 0 \cdots 0 \\ \vdots \quad \vdots & \vdots & \vdots \quad \vdots \\ 0 \cdots 0 & f_n^{(r)} & 0 \cdots 0 \end{bmatrix} \tag{4}$$

This is the matrix whose columns are all zero except the sth, which is $\mathbf{f}^{(r)}$ Then, since the sth entry in the vector $\mathbf{y}^{(r)}$ is 1, we have

$$Z^{(r)}\mathbf{y}^{(r)} = \mathbf{f}^{(r)},$$

and from Equations (2) and (3) we see that the matrix $Z^{(r)}$ of (4) satisfies our requirement.

We can now estimate the accuracy of the eigensolution computed by the iterative method, in which we terminate the process when two successive vectors $\mathbf{y}^{(r)}$ and $\mathbf{y}^{(r+1)}$ are the same to the number of figures to which we we are working. By (3), the resulting vector $\mathbf{y}^{(r+1)}$ and the number $k^{(r+1)}$ constitute an eigensolution of $A + Z^{(r)}$, and we can then apply the theory of Section 30.2. For the eigenvalue, Equation (5) of sub-section 30.2.1 gives the result

$$\begin{aligned} |k^{(r+1)} - \text{true eigenvalue}| &\leqslant M(Z^{(r)}) \\ &= M(\mathbf{f}^{(r)}), \end{aligned}$$

from Equation (4), and this is the magnitude of the entry of $\mathbf{f}^{(r)}$ of largest magnitude.

Now if no entry of A exceeds unity in absolute value, and since no entry of $\mathbf{y}^{(r)}$ exceeds unity in absolute value because of (1), we see that the entries of $\mathbf{f}^{(r)}$ can hardly exceed

$$n \times 2^{-t},$$

where t is the number of binary digits to which our computer performs its arithmetic. Indeed this is guaranteed if we exercise some slight extra care with the arithmetic, as described on page 35 of *Unit 8*.

The computed vector, of course, could be far less accurate, with a possible error of

$$\frac{n \times 2^{-t}}{\lambda_1 - \lambda_s}$$

in its components, and this could be large if any λ_s is near to λ_1. But this merely reflects the possible inherent instability of the eigenvectors, and we can hardly blame our *method* for this!

Induced instability will occur, however, if we try to find any other solution by direct iteration starting with a vector $\mathbf{y}^{(0)}$ which is orthogonal to the dominant eigenvector, that is with

$$\mathbf{y}^{(0)} = \alpha_2 \mathbf{x}_2 + \alpha_3 \mathbf{x}_3 + \cdots + \alpha_n \mathbf{x}_n$$

(assuming that λ_1 is the eigenvalue of largest magnitude).

If we have computed λ_1 and $\mathbf{x}_1$ *exactly*, then we can find such a $\mathbf{y}^{(0)}$ by the Gram-Schmidt method described in *Unit 16, Euclidean Spaces I*, starting with some arbitrary vector $\mathbf{y}$, and taking

$$\mathbf{y}^{(0)} = \mathbf{y} - \alpha_1 \mathbf{x}_1$$

with

$$\alpha_1 = \mathbf{x}_1^T\mathbf{y} \text{ if } \mathbf{x}_1^T\mathbf{x}_1 = 1.$$

In theory direct iteration with A will now converge to a different eigensolution; but in practice the method will fail because (a) we will not have $\mathbf{x}_1$ exactly, (b) we cannot compute α_1 or $\mathbf{y}^{(0)}$ exactly, and (c) even if we could, the rounding errors in the iterative process, which worked to our advantage in computing the dominant eigensolutions, now magnify as the calculation proceeds, and insist on producing once again one of these dominant solutions.

Exercise

For the matrix $\begin{bmatrix} 1 & 0.01 \\ 0.01 & 1 \end{bmatrix}$ the dominant eigensolution is

$$\lambda_1 = 1.01, \mathbf{x}_1 = [1 \quad 1]^T$$

scaled so that $M(\mathbf{x}_1) = 1$.

Using direct iteration with the starting vector

$$\mathbf{y}^{(0)} = [0.7 \quad 1]^T$$

and using one-digit arithmetic, show that we obtain an accurate eigenvalue but an inaccurate eigenvector. Explain this result.

Solution

The iteration produces

r	$\mathbf{z}^{(r)T}$	$\mathbf{y}^{(r)T}$	λ
0	[0.7 1]	[0.7 1]	
1	[0.7 1.0]	[0.7 1.0]	1.0

The method converges rapidly (in one step!) and the computed eigenvalue is very accurate, but the computed vector is rather inaccurate. This occurs because the eigenvalues are 1.01 and 0.99, very close together, and the rounding errors of the arithmetic are equivalent to the solution of a slightly perturbed matrix for which (both) eigenvectors may be quite different from those of the original matrix.

30.3.4 Inverse Iteration

We now discuss a method of iteration which will give any required eigensolution. A hint of the idea is given by the title *inverse iteration.* In fact, inverse iteration with the matrix $A - pI$, where p is any number, is effectively direct iteration with its inverse $(A - pI)^{-1}$.

For the theory of inverse iteration we need the eigenvectors and eigenvalues of $A - pI$ in terms of those of A. We start from the formula

$$(A - pI)\mathbf{x}_r = (\lambda_r - p)\mathbf{x}_r$$

which we have already used for direct iteration with $A - pI$. Provided $\lambda_r \neq p$, the matrix $A - pI$ is non-singular, and we can therefore premultiply both sides by $(\lambda_r - p)^{-1}(A - pI)^{-1}$ and obtain

$$(\lambda_r - p)^{-1}\mathbf{x}_r = (A - pI)^{-1}\mathbf{x}_r.$$

That is, if $\lambda_r \neq p$, then $\mathbf{x}_r$ is an eigenvector of $(A - pI)^{-1}$ corresponding to the eigenvalue $(\lambda_r - p)^{-1}$.

The effect of operating on any vector

$$\mathbf{y} = \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n$$

repeatedly with $(A - pI)^{-1}$ is then given by the formula

$$(A - pI)^{-r}\mathbf{y} = \alpha_1(\lambda_1 - p)^{-r}\mathbf{x}_1 + \cdots + \alpha_n(\lambda_n - p)^{-r}\mathbf{x}_n,$$

which has obvious analogies with Formula (3) of sub-section 30.3.1. As r increases, the term in the sum that increases fastest is the one for which $|(\lambda_k - p)^{-1}|$ is the largest. This, of course, is the term for which $|\lambda_k - p|$ is the *smallest*. As r increases, therefore, the right-hand side gets closer and closer to a multiple of the eigenvector $\mathbf{x}_k$ whose eigenvalue is closest to p.

The corresponding iterative algorithm, again, can be written in a form analogous to that of Equations (1) of sub-section 30.3.2 relevant to direct iteration, and would satisfy the equations:

$$\begin{aligned} &\mathbf{y}^{(0)} \text{ arbitrary (with largest entry} = 1) \\ &\mathbf{z}^{(r+1)} = (A - pI)^{-1}\mathbf{y}^{(r)} \qquad (1) \\ &\mathbf{y}^{(r+1)} = \mathbf{z}^{(r+1)}/m(\mathbf{z}^{(r+1)}) \end{aligned}$$

Then $\mathbf{y}^{(r)}$ gets nearer to an eigenvector $\mathbf{x}_k$ as r increases, the corresponding eigenvalue λ_k is the one nearest to p, and its computation comes from the equation

$$m(\mathbf{z}^{(r+1)}) \simeq (\lambda_k - p)^{-1},$$

the approximation getting steadily more accurate with increasing r.

The choice $p = 0$, for example, produces the eigenvalue of smallest magnitude, and this is often required in practical problems. We can also solve immediately the problem of finding the eigenvalue nearest to a particular *given* value of p. It is in fact clear that inverse iteration, which can find any solution with suitable choice of p, is far more powerful than direct iteration which, whatever choice we make for p, can only produce the algebraically largest and smallest eigenvalues.

In practice we change the algorithm (1) in one important way. It would appear that we have to invert the matrix $(A - pI)$, and we have seen in *Unit 8* that this is a somewhat laborious operation. It is quicker to solve linear equations, and we therefore write the second of Equations (1) in the form

$$(A - pI)\mathbf{z}^{(r+1)} = \mathbf{y}^{(r)}$$

which is theoretically the same but computationally quite different!

Notice, in particular, that for each r the linear equations have the *same matrix* $A - pI$ but *different right-hand sides*. In *Unit 8* we saw that the largest part of the Gaussian elimination process with interchanges is the reduction of the relevant row-permutation of $A - pI$ to an upper triangle U, and this need be done once only. To calculate each new $\mathbf{z}^{(r+1)}$ we use the same multipliers and interchanges on $\mathbf{y}^{(r)}$ as we used in reducing $A - pI$ to U. This converts $\mathbf{y}^{(r)}$ to $\mathbf{c}^{(r)}$ such that

$$U\mathbf{z}^{(r+1)} = \mathbf{c}^{(r)},$$

a system of equations from which $\mathbf{z}^{(r+1)}$ is found by back substitution. This is far more economical than *computing* the inverse $(A - pI)^{-1}$ and then evaluating $(A - pI)^{-1}\mathbf{y}^{(r)}$ to get $\mathbf{z}^{(r+1)}$.

Example

Let us iterate inversely with the matrix

$$A = \begin{bmatrix} 2 & 1 & 4 \\ 1 & -1 & -4 \\ 4 & -4 & -8 \end{bmatrix}.$$

The first step is to perform the relevant triangular decomposition, and since in *Unit 8* we only sketched the decomposition process corresponding to interchanges in Gaussian elimination we will go through it here in some detail. We proceed to perform the elimination with interchanges, but keeping close account of the *positions* of the *original* rows in A. The numbers in brackets refer to the rows of A, and we start with

$$\begin{array}{rrrl} 2 & 1 & 4 & (1) \\ 1 & -1 & -4 & (2) \\ 4 & -4 & -8 & (3). \end{array}$$

We interchange rows (1) and (3) and then perform the elimination. The multipliers are written on the left, and for simplicity we use exact arithmetic.

$$\begin{array}{rrrrl} & 4 & -4 & -8 & (3) \\ -\frac{1}{4} & 1 & -1 & -4 & (2) \\ -\frac{1}{2} & 2 & 1 & 4 & (1) \end{array}$$

$$\begin{array}{rrrl} 4 & -4 & -8 & (3) \\ 0 & 0 & -2 & (2) \\ 0 & 3 & 8 & (1) \end{array}$$

We now interchange the second and third rows of *this* matrix, and then eliminate, obtaining

$$\begin{array}{rrrrl} & 4 & -4 & -8 & (3) \\ & 0 & 3 & 8 & (1) \\ 0 & 0 & 0 & -2 & (2). \end{array}$$

This shows that the upper triangle U comes from $\bar{A}$, which is A with rows written in order 3, 1, 2. We know that the non-diagonal elements of the L matrix (in $\bar{A} = LU$) are the negatives of the multipliers, but what are their positions in the L matrix? Well, we just look for the multipliers associated with the rows of $\bar{A}$. The first is the third row of A, and this has no multipliers. The second is the first row of A, and this has the single multiplier $-\frac{1}{2}$. The third is the second row of A, and this has multipliers $-\frac{1}{4}$ and 0 *in this order*. The required L matrix is therefore

$$\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{4} & 0 & 1 \end{bmatrix}$$

and we easily verify that

$$\overset{\bar{A}}{\begin{bmatrix} 4 & -4 & -8 \\ 2 & 1 & 4 \\ 1 & -1 & -4 \end{bmatrix}} = \overset{L}{\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{4} & 0 & 1 \end{bmatrix}} \overset{U}{\begin{bmatrix} 4 & -4 & -8 \\ 0 & 3 & 8 \\ 0 & 0 & -2 \end{bmatrix}}$$

Now we start the inverse iteration, taking the arbitrary $\mathbf{y}^{(0)}$ as

$$\mathbf{y}^{(0)T} = [0.32 \quad 1.0 \quad -0.32].$$

The $\mathbf{c}^{(0)}$ corresponding to $\mathbf{y}^{(0)}$ is obtained by subjecting $\mathbf{y}^{(0)}$ to the same row interchanges and multipliers that took A to $\bar{A}$ and thence to U. In other words,

$$L\mathbf{c}^{(0)} = [-0.32 \quad 0.32 \quad 1.0]^T,$$

and solving these linear equations by *forward* substitution we find

$$\mathbf{c}^{(0)} = [-0.32 \quad 0.48 \quad 1.08]^T.$$

Our $\mathbf{z}^{(1)}$ is then obtained by *back* substitution in the equations

$$U\mathbf{z}^{(1)} = \mathbf{c}^{(0)},$$

and we easily find

$$z_3^{(1)} = -0.54, z_2^{(1)} = 1.60, z_1^{(1)} = 0.44.$$

Then

$$\mathbf{y}^{(1)T} = [0.275 \quad 1 \quad -0.3375]$$

and the current approximation to the eigenvalue is obtained from

$$m(z^{(1)}) \simeq \lambda^{-1} \qquad (\text{since } p = 0),$$

giving

$$\lambda \simeq 1/1.60 = 0.625.$$

Compared with direct iteration, the only extra feature of the inverse iteration method which could possibly give rise to induced instability is the solution of the relevant linear equations. This we know to be a stable process if we use Gauss elimination with interchanges. But however stable the method, we may have an ill-conditioned mathematical problem if the selected p, in inverse iteration with $A - pI$, is very nearly an eigenvalue of A. For then the matrix $A - pI$ is almost singular, and as we saw in *Unit 8* this is just the situation in which ill-conditioning is most manifest.

It is extremely unlikely, of course, that our selected p will be very near to an eigenvalue, and even then, as we shall show in sub-section 30.4.3, there is no trouble if the vector $\mathbf{y}^{(r)}$, say, in Equation (1), is not nearly orthogonal to the eigenvector $\mathbf{x}_k$ corresponding to the eigenvalue λ_k which is near to p. As we saw with direct iteration, for any significant value of r, that is after a fair number of steps, the vector $\mathbf{y}^{(r)}$ will have picked up a good component of $\mathbf{x}_k$, if only through the magnification of earlier rounding errors, and we have no induced instability.

Exercise

The eigenvalues of the matrix $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 3 & 1 & 4 \end{bmatrix}$ in the example of subsection 30.3.2 are approximately 6.4, -1.3 and -0.1. If we iterate inversely with $A - pI$, with $p = 4$, to what eigensolution shall we converge, and what are the main factors governing the rate of convergence?

Solution

The number $p = 4$ is closest to the eigenvalue 6.4, so that we shall converge to the dominant eigensolution. We have

$$6.4 - p = 2.4, \; -1.3 - p = -5.3, \; -0.1 - p = -4.1,$$

so that convergence depends on the rates at which

$$\left(\frac{2.4}{5.3}\right)^r \quad \text{and} \quad \left(\frac{2.4}{4.1}\right)^r$$

get small. Clearly the second one is the more important.

30.3.5 Summary of Section 30.3

In this section we defined the terms

direct iteration with A and with $A - pI$	(page C20)	* *
inverse iteration with A and with $A - pI$	(page C27)	* *

We introduced the notation

$m(\mathbf{z})$	(page C22)

Techniques

1. Use direct iteration, with the matrix $A - pI$, to converge to one of the extreme eigensolutions. *
2. Find a p which gives most rapid convergence to one of these eigensolutions. *
3. Use inverse iteration with A to converge to the eigensolution corresponding to the eigenvalue of smallest magnitude. * *
4. Use inverse iteration with $A - pI$ to converge to the eigensolution whose eigenvalue is closest to a given p. * *
5. Use stable Gauss elimination with interchanges for solving linear equations in inverse iteration. *

30.4 THE METHOD OF SIMILARITY TRANSFORMATIONS

30.4.0 Introduction

We turn now to a type of method which is suitable when we want all or many of the eigensolutions of A instead of just a few. The basis of all such methods is the determination of a non-singular matrix P such that the eigensolutions of the matrix $B = P^{-1}AP$ can be found quickly and easily. A and B are similar matrices and the transformation from A to B is a *similarity transformation* (of matrices) and therefore preserves eigenvalues. Moreover, we can easily find the eigenvectors of A from those of B, since if $\mathbf{y}_r$ is an eigenvector of B, then

$$B\mathbf{y}_r = \lambda_r \mathbf{y}_r,$$

so that

$$A(P\mathbf{y}_r) = P(P^{-1}AP)\mathbf{y}_r = PB\mathbf{y}_r = \lambda_r(P\mathbf{y}_r)$$

and $P\mathbf{y}_r$ is an eigenvector of A.

In using the method we do not try to find the matrix of transition P at one step; we perform a succession of similarity transformations which simplify the given matrix by stages.

When the matrix A is symmetric we want our transformations to preserve the symmetry. This is achieved by taking P to be an orthogonal matrix, so that

$$P^{-1} = P^T.$$

Then

$$B = P^TAP,$$

which is also symmetric. In the method we shall describe, each step in the sequence of transformations corresponds to a rotation in the plane of one pair of coordinate axes. It is possible, with an infinite sequence of such transformations (which of course in practice is terminated at some suitable point!), to transform A to a diagonal matrix D, whose eigenvalues are just its diagonal elements. It is more convenient, however, to use a *finite* number of steps which transform A to a *triple-diagonal matrix*, with non-zero entries only on the main diagonal and on each of the adjacent sloping lines. It is fairly easy to determine the eigensolutions of a triple-diagonal matrix. In the next sub-section we show how to obtain the triple-diagonal matrix, and after that we discuss a technique for finding one or more of its eigensolutions.

30.4.1 Similarity Transformation to Triple-diagonal Form

We propose to reduce the matrix A to a simpler matrix B by a systematic replacement by zeros of the elements of A outside the main diagonal and the two adjacent sloping lines. As an example of how the procedure works we will look at a 4 × 4 case. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{12} & a_{22} & a_{23} & a_{24} \\ a_{13} & a_{23} & a_{33} & a_{34} \\ a_{14} & a_{24} & a_{34} & a_{44} \end{bmatrix}, \quad P_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_1 & -\sin\theta_1 & 0 \\ 0 & \sin\theta_1 & \cos\theta_1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

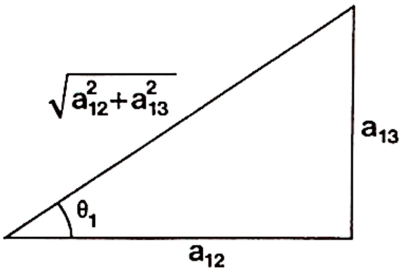
This orthogonal matrix P_1 represents a rotation in 4-space through an angle θ about the origin in the (2, 3) plane. We shall choose the angle θ_1 to make a pair of elements in the transformed matrix $A_1 = P_1^TAP_1$ equal to 0. Carrying out the matrix multiplication, we find that the (1, 3) and (3, 1) entries of A_1 are

$$a_{13}^{(1)} = a_{31}^{(1)} = -s_1 a_{12} + c_1 a_{13},$$

where $s_1 = \sin \theta_1$, $c_1 = \cos \theta_1$.

We can make these entries zero by taking

$$c_1 = \frac{a_{12}}{(a_{12}^2 + a_{13}^2)^{1/2}}, \quad s_1 = \frac{a_{13}}{(a_{12}^2 + a_{13}^2)^{1/2}}.$$



The matrix A_1, as you can easily verify, looks like

$$A_1 = \begin{bmatrix} a_{11} & a_{12}^{(1)} & 0 & a_{14} \\ a_{12}^{(1)} & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ 0 & a_{23}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} \\ a_{14} & a_{24}^{(1)} & a_{34}^{(1)} & a_{44} \end{bmatrix}$$

all the entries in rows and columns 2 and 3 having been changed by the transformation.

The next step is to make a transformation which reduces another pair of entries to zero, while leaving untouched the zeros we have already produced in the (1, 3) and (3, 1) positions of A_1. This can be accomplished by a rotation in the (2, 4) plane, which will change the entries only in the second and fourth rows and columns. The new matrix of transition is

$$P_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c_2 & 0 & -s_2 \\ 0 & 0 & 1 & 0 \\ 0 & s_2 & 0 & c_2 \end{bmatrix}$$

and the new symmetric matrix is

$$A_2 = P_2^T A_1 P_2.$$

You can verify, by carrying out the matrix multiplications, that the (1, 4) and (4, 1) entries of A_2 will be zero if we take

$$c_2 = \frac{a_{12}^{(1)}}{[(a_{12}^{(1)})^2 + (a_{14})^2]^{1/2}}, \quad s_2 = \frac{a_{14}}{[(a_{12}^{(1)})^2 + (a_{14})^2]^{1/2}}$$

The new matrix has the appearance

$$A_2 = \begin{bmatrix} a_{11} & a_{12}^{(2)} & 0 & 0 \\ a_{12}^{(2)} & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ 0 & a_{23}^{(2)} & a_{33}^{(1)} & a_{34}^{(2)} \\ 0 & a_{24}^{(2)} & a_{34}^{(2)} & a_{44}^{(2)} \end{bmatrix}$$

We have now finished with the first row and column, and proceed to consider the reduction to zero of the entries in the (2, 4) and (4, 2) positions of A, This is achieved by the transformation

$$A_3 = P_3^T A_2 P_3,$$

where P_3^T is a rotation matrix in the (3, 4) plane, with

$$c_3 = \frac{a_{23}^{(2)}}{[(a_{23}^{(2)})^2 + (a_{24}^{(2)})^2]^{1/2}}, \quad s_3 = \frac{a_{24}^{(2)}}{[(a_{23}^{(2)})^2 + (a_{24}^{(2)})^2]^{1/2}}$$

The entries which change are those in rows and columns 3 and 4, but the zeros in the first row and column do not change because the new entries in these positions are just linear combinations of the old entries, and combinations of zeros produce zeros!

For our matrix of order 4 we have now finished, the triple-diagonal form looking like

$$C = \begin{bmatrix} a_1 & b_2 & 0 & 0 \\ b_2 & a_2 & b_3 & 0 \\ 0 & b_3 & a_3 & b_4 \\ 0 & 0 & b_4 & a_4 \end{bmatrix}$$

a notation which we shall find useful in the next sub-section.

Example

$$\text{Let } A = \begin{bmatrix} 12 & 3 & 4 & 12 \\ 3 & -12 & 0 & 3 \\ 4 & 0 & -12 & 4 \\ 12 & 3 & 4 & 12 \end{bmatrix}$$

We will reduce A to triple-diagonal form. At the first stage $a_{12} = 3$, $a_{13} = 4$, so that $c_1 = \frac{3}{5}$, $s_1 = \frac{4}{5}$ and we have

$$P_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{3}{5} & -\frac{4}{5} & 0 \\ 0 & \frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

We calculate $A_1 = P_1^T A P_1$, and obtain the matrix

$$A_1 = \begin{bmatrix} 12 & 5 & 0 & 12 \\ 5 & -12 & 0 & 5 \\ 0 & 0 & -12 & 0 \\ 12 & 5 & 0 & 12 \end{bmatrix}$$

At the next stage we look at the (1, 4) and (4, 1) entries of A_1. $a_{12}^{(1)} = 5$, $a_{14}^{(1)} = 12$, so that

$$c_2 = \tfrac{5}{13}, \quad s_2 = \tfrac{12}{13}$$

and we take

$$P_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{5}{13} & 0 & -\frac{12}{13} \\ 0 & 0 & 1 & 0 \\ 0 & \frac{12}{13} & 0 & \frac{5}{13} \end{bmatrix}$$

We calculate $A_2 = P_2^T A_1 P_2$ to find

$$A_2 = \begin{bmatrix} 12 & 13 & 0 & 0 \\ 13 & 12 & 0 & 5 \\ 0 & 0 & -12 & 0 \\ 0 & 5 & 0 & -12 \end{bmatrix}$$

At the last stage we reduce the entry 5 in the (2, 4), (4, 2) positions to zero. $a_{23}^{(2)} = 0$, $a_{24}^{(2)} = 5$, so that

$$c_3 = 0, \quad s_3 = 1$$

and we take

$$P_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Performing the appropriate matrix multiplications we calculate $A_3 = P_3^T A_2 P_3$ and obtain

$$A_3 = \begin{bmatrix} 12 & 13 & 0 & 0 \\ 13 & 12 & 5 & 0 \\ 0 & 5 & -12 & 0 \\ 0 & 0 & 0 & -12 \end{bmatrix}$$

which is in triple-diagonal form. Notice that this is rather a special triple-diagonal form, two of the entries in the sloping lines next to the diagonal being zeros. Clearly $\lambda = -12$ is an eigenvalue of A_3, and the others are the eigenvalues of the leading 3×3-submatrix of A_3, which is in "normal" triple-diagonal form.

For the general case we produce the required zeros in the first row and column by successive rotations in the planes (2, 3), (2, 4), ..., (2, n). Zeros in the second row and column are produced by rotations in planes (3, 4), (3, 5), ..., (3, n); and so on. The number of successive transformations is

$$(n-2) + (n-3) + \cdots + 1 = \tfrac{1}{2}(n-1)(n-2),$$

and we have performed the operation

$$P_N^T P_{N-1}^T \cdots P_2^T P_1^T A P_1 P_2 \cdots P_{N-1} P_N = C,$$

where C is a triple-diagonal matrix and

$$N = \tfrac{1}{2}(n-1)(n-2).$$

An eigenvalue of C is an eigenvalue of A, and if $\mathbf{y}$ is an eigenvector of C, the corresponding eigenvector of A is

$$\mathbf{x} = (P_1 P_2 \cdots P_N)\mathbf{y}.$$

When we have computed an eigenvector $\mathbf{y}$ of C we compute the corresponding $\mathbf{x}$ by successively forming $P_N \mathbf{y}$, $P_{N-1}(P_N \mathbf{y}), \ldots, P_1(P_2 \ldots P_N)\mathbf{y}$, rather than by first computing the matrix product $(P_1 P_2 \cdots P_N)$.

Exercise

Reduce the matrix $A = \begin{bmatrix} 2 & 4 & -3 \\ 4 & 2 & -3 \\ -3 & -3 & -6 \end{bmatrix}$

to a triple-diagonal form.

Solution

At the first stage we reduce to zero the entries in the (1, 3) and (3, 1) positions. We take

$$c_1 = \frac{4}{(4^2 + 3^2)^{1/2}} = \frac{4}{5}, \quad s_1 = -\frac{3}{(4^2 + 3^2)^{1/2}} = -\frac{3}{5}$$

and

$$P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{4}{5} & \frac{3}{5} \\ 0 & -\frac{3}{5} & \frac{4}{5} \end{bmatrix}.$$

Then

$$A_1 = P_1^T A P_1 = \begin{bmatrix} 2 & 5 & 0 \\ 5 & 2 & 3 \\ 0 & 3 & -6 \end{bmatrix}$$

which is already in triple-diagonal form.

30.4.2 Eigenvalues of a Symmetric Triple-diagonal Matrix

We could use any of our iterative methods to find eigensolutions of a triple-diagonal matrix, and the large number of zero entries in this matrix considerably reduces the amount of computation involved. But symmetric matrices have some special properties which give us a much more attractive method, at least for the determination of the eigenvalues. Consider the symmetric matrix

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 3 & 1 & 4 \end{bmatrix}$$

It has *leading* or *principal submatrices*

$$[2] \quad \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 3 & 1 & 4 \end{bmatrix}$$

The eigenvalues of these submatrices are approximately

$$2, \quad \{2.303, -1.303\}, \quad \{6.425, -1.306, -0.119\}$$

respectively.

Inspection of these eigenvalues shows that the eigenvalue of [2] separates those of $\begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$:

$$-1.303 < 2 < 2.303;$$

and the eigenvalues of $\begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$ separate those of

$$\begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 3 & 1 & 4 \end{bmatrix};$$

$$-1.306 < -1.303 < -0.119 < 2.303 < 6.425$$

This result generalizes to give the first important property of symmetric matrices; we state it as a theorem but do not prove it.

Theorem 1

For a symmetric matrix, the eigenvalues of the principal submatrix of order r *separate* those of the principal submatrix of order $r + 1$.

This result leads to a second theorem, which we àgain state without proof, but which is the basis of a very practical method for computing the eigenvalues of a triple-diagonal matrix.

Theorem 2

If we take a particular value of p in the matrix $A - pI$, and compute the determinants of its principal submatrices, calling these values $f_1, f_2, f_3, \ldots, f_n$ for a matrix of order n, then the number of agreements in sign between successive members of the sequence $f_0, f_1, \ldots, f_n$, where $f_0 = 1$, is equal to the number of eigenvalues of the matrix A greater than the number p.

Examples

(i) With $p = 0$, the successive principal determinants of the matrix

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 3 & 1 & 4 \end{bmatrix}$$

are

f_0	f_1	f_2	f_3
1	2	-3	1

There is one agreement in sign between successive members, between f_0 and f_1, and therefore A has just one positive eigenvalue (since $p = 0$).

(ii) With $p = -1$, we find for the matrix $A + I$ the sequence

f_0	f_1	f_2	f_3
1	3	-1	-2

Here there are two agreements of sign, which tells us that A has two eigenvalues greater than -1. Since there is only one positive eigenvalue, as we have shown, it follows that there is just one eigenvalue between -1 and 0.

The point of Theorem 2 is that it enables us quite quickly to isolate any particular root within a particular interval. To locate it more closely we would now proceed to use a *process of bisection*, taking next $p = -0.5$ in our example. If in the resulting sequence there are two agreements in sign, we deduce that there is one eigenvalue between -0.5 and 0 (and one greater than 0); and if there is only one agreement, we know that there is one eigenvalue in the interval $(-1, -0.5)$.

Once we have located an eigenvalue in an interval (a, b), then after k bisections we have confined it to an interval of length $2^{-k}(b - a)$, and convergence to an accurate result is quite rapid. On the other hand we have many determinants to evaluate, and this is where the triple-diagonal form C shows to real advantage.

What does the trick, as you might expect, is a recurrence relation! For a 4×4-matrix, using our previous notation, we have

$$C - pI = \begin{bmatrix} a_1 - p & b_2 & 0 & 0 \\ b_2 & a_2 - p & b_3 & 0 \\ 0 & b_3 & a_3 - p & b_4 \\ 0 & 0 & b_4 & a_4 - p \end{bmatrix}$$

The determinants of the principal submatrices are

$$f_1 = a_1 - p,$$
$$f_2 = (a_2 - p)(a_1 - p) - b_2^2 = (a_2 - p)f_1 - b_2^2 f_0$$

where $f_0 = 1$. Continuing in this way it can be shown, for a general triple-diagonal matrix, that

$$f_{r+1} = (a_{r+1} - p)f_r - b_{r+1}^2 f_{r-1} \qquad (1)$$

a second-order recurrence relation.

Example

Our orthogonal similarity transformation applied to the matrix

$$A = \begin{bmatrix} 1 & \sqrt{2} & \sqrt{2} & 2 \\ \sqrt{2} & -\sqrt{2} & -1 & \sqrt{2} \\ \sqrt{2} & -1 & \sqrt{2} & \sqrt{2} \\ 2 & \sqrt{2} & \sqrt{2} & -3 \end{bmatrix}$$

produces, with exact arithmetic, the triple-diagonal form

$$C = \begin{bmatrix} 1 & -2\sqrt{2} & 0 & 0 \\ -2\sqrt{2} & 0 & \sqrt{2} & 0 \\ 0 & \sqrt{2} & -\frac{1}{2} & -\frac{5}{2} \\ 0 & 0 & -\frac{5}{2} & -\frac{5}{2} \end{bmatrix}$$

If the eigenvalues are $\lambda_1, \lambda_2, \lambda_3$ and λ_4 in decreasing order, let us find an interval which contains *only* λ_3.

We saw in sub-section 30.1.1 that no eigenvalue exceeds $M(C)$ in magnitude. Here $M(C) = 5$ (from its last row), so that all eigenvalues are in the interval $[-5, 5]$. Let us bisect this and take $p=0$. Using the recurrence relation, we find

f_0	f_1	f_2	f_3	f_4
1	1	-8	2	45

There are two agreements in sign and therefore two positive eigenvalues, which must be λ_1 and λ_2. The required λ_3 is therefore in the interval $[-5, 0)$, as also is λ_4. Next we bisect this interval, take $p = -2.5$, and compute the sequence

f_0	f_1	f_2	f_3	f_4
1	3.5	0.75	-5.5	-4.6875

Three agreements in sign mean that λ_3 is in the interval $(-2.5, 0]$, and λ_4 is in the interval $[-5, -2.5]$.

Either of these eigenvalues can be located more accurately by successive bisections of the relevant intervals.

Exercise

The triple-diagonal matrix

$$C = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

is positive definite. Find an interval, not exceeding unity in length, which contains the smallest eigenvalue. (If any $f_r = 0$, its "sign" should be taken as opposite to that of f_{r-1}.)

Solution

Since C is positive definite all its eigenvalues are positive, by Equation (5) of sub-section 30.1.1. Also $M(C) = 4$; so all eigenvalues lie in $(0, 4]$. We start with $p = 2$, and produce the sequence

f_0	f_1	f_2	f_3	f_4
1	0	-1	0	1

Noting the comment about the "sign" of zero, we find signs

$$+ \quad - \quad - \quad + \quad +$$

and the existence of two agreements shows that the two smallest roots lie in $(0, 2]$.

Bisecting $(0, 2)$, and taking $p = 1$, we get

f_0	f_1	f_2	f_3	f_4
1	1	0	-1	-1

Here there are three agreements in sign, and our required smallest root lies in the interval $(0, 1]$.

30.4.3 Eigenvectors of a Triple-diagonal Matrix

Once we have found an eigenvalue we can in theory compute the corresponding eigenvector by solving the linear homogeneous equations

$$(C - \lambda I)\mathbf{x} = \mathbf{0}.$$

For a matrix of order 3, for example, this system has the form

$$\begin{aligned} (a_1 - \lambda)x_1 + \quad b_2 x_2 \qquad\qquad &= 0 \\ b_2 x_1 + (a_2 - \lambda)x_2 + \quad b_3 x_3 &= 0 \\ b_3 x_2 + (a_3 - \lambda)x_3 &= 0 \cdot \end{aligned} \tag{1}$$

where $\mathbf{x} = [x_1\ x_2\ x_3]^T$.

Let us first consider the exact solution of a system such as (1) assuming that our computed eigenvalue is an *exact* eigenvalue. We assume that b_2, b_3 are both non-zero and look for a solution scaled so that $x_1 = 1$. Then we can calculate x_2 from the first equation and x_3 from the second, and the last will then be satisfied automatically since the matrix $C - \lambda I$ is singular, so that its rows are linearly dependent. It turns out that x_1, x_2, x_3 can be expressed in terms of the leading determinants f_0, f_1, f_2, for our method gives

$$\begin{aligned} x_1 &= 1 = f_0 \\ x_2 &= -(a_1 - \lambda)/b_2 = -f_1/b_2 \\ x_3 &= -(b_2 x_1 + (a_2 - \lambda)x_2)/b_3 \\ &= -\left(b_2 f_0 - (a_2 - \lambda)\frac{f_1}{b_2}\right)/b_3 \\ &= f_2/(b_2 b_3) \end{aligned} \tag{2}$$

by the recurrence relation (1) of the preceding sub-section. For an $n \times n$-triple-diagonal matrix, we obtain by the same method the result

$$x_r = (-1)^{r-1} f_{r-1}/(b_2 b_3 \ldots b_r) \qquad (r = 1, \ldots, n) \tag{3}$$

Although correct *in theory*, the method we have just described turns out to be quite unsatisfactory for practical computations with inexact arithmetic; it suffers from acute induced instability. An example will reveal the dangers of this approach.

Example

The 21 × 21-triple-diagonal matrix

$$C = \begin{bmatrix} 10 & 1 & & & & & \\ 1 & 9 & 1 & & & & \\ & 1 & 8 & 1 & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & \cdot & \cdot & \cdot \\ & & & & 1 & -9 & 1 \\ & & & & & 1 & -10 \end{bmatrix}$$

has an eigenvalue $\lambda \simeq 10.746\,194\,183$, correct to 11 significant figures. If we take the extremely good approximation $\lambda = 10.746\,194\,2$, correct to nine figures, and denote the corresponding eigenvector by $[x_1 \ldots x_{21}]^T$, we can take $x_1 = 1$ and compute the numbers $x_2, x_3, \ldots, x_{21}$ in succession from the first 20 equations in the system $(C - \lambda I)\mathbf{x} = \mathbf{0}$. The result gives a vector some of whose entries are approximately

x_1	x_4	x_7	x_{11}	x_{15}	x_{21}
1	0.09	0.0005	0.036	0.77×10^3	0.19×10^{11}

The corresponding entries of the correct eigenvector are completely different; they are:

$$1 \quad 0.09 \quad 0.0005 \quad 7 \times 10^{-8} \quad 2 \times 10^{-12} \quad 0.7 \times 10^{-19}$$

and we have a catastrophic case of induced instability! [This example is taken from: J. H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon Press, Oxford 1965).]

We can explain this result using backward error analysis, by finding the equations which our computed "eigenvector" satisfy exactly. Assuming no mistakes in the arithmetic, other than a small error in our approximate eigenvalue, which we will call $\bar{\lambda}$, we have satisfied *exactly* the first $n-1$ of the equations $(C - \lambda I)\mathbf{x} = \mathbf{0}$. The last equation is obviously not satisfied, since otherwise both the assumed eigenvalue and the computed eigenvector would be exact. Suppose that substitution in the last equation produces the number m. Then we have satisfied exactly the equations

$$(C - \bar{\lambda} I)\mathbf{x} = m\mathbf{e}_n \tag{4}$$

where $\mathbf{e}_n$ is the last column of the unit matrix. What is the solution of these equations? Writing

$$\mathbf{x} = \sum_{r=1}^{n} \alpha_r \mathbf{x}_r \tag{5}$$

where $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are the normalized eigenvectors of C, with eigenvalues $\lambda_1, \ldots, \lambda_n$, we find that (4) becomes

$$\sum_{r=1}^{n} \alpha_r(\lambda_r - \bar{\lambda})\mathbf{x}_r = m\mathbf{e}_n$$

Taking the inner product with any eigenvector $\mathbf{x}_s$ and using

$$\mathbf{x}_r \cdot \mathbf{x}_s = \mathbf{x}_r^T \mathbf{x}_s = \delta_{rs},$$

we obtain

$$(\lambda_s - \bar{\lambda})\alpha_s = m\mathbf{e}_n \cdot \mathbf{x}_s \qquad (s = 1, \ldots, n)$$

so that (5) (with s used in place of r) becomes

$$\mathbf{x} = \sum_{s=1}^{n} \frac{m\mathbf{e}_n \cdot \mathbf{x}_s}{\lambda_s - \bar{\lambda}} \mathbf{x}_s \tag{6}$$

If $\bar{\lambda}$ is very close to a particular eigenvalue λ_k, and $m\mathbf{e}_n \cdot \mathbf{x}_k$ is not very small, then the term with $s = k$ in (6) will dominate, and our calculated vector $\mathbf{x}$ is very nearly a multiple of the eigenvector $\mathbf{x}_k$ which we are trying to compute. In our particular example, however, $m\mathbf{e}_n \cdot \mathbf{x}_k$ *is* very small, since the nth entry in the vector $\mathbf{x}_k$ is 0.7×10^{-19}. Our computed "eigenvector" is therefore far from being proportional to $\mathbf{x}_k$. Such a possibility might appear to be unlikely to occur in practice, but experience shows that it is in fact quite common with triple-diagonal matrices obtained (from symmetric matrices) by orthogonal similarity transformations.

Now we notice, from Equation (4), that out process is equivalent to one step of *inverse iteration* with a particular starting vector $\mathbf{y}^{(0)} = m\mathbf{e}_n$, and it has failed because the matrix $C - \bar{\lambda} I$ is nearly singular and because the starting vector is nearly orthogonal to the required eigenvector. Our analysis shows that these two causes act together to produce the failure. If we remove one of them, by using a starting vector which is *not* nearly orthogonal to the required eigenvector, then the process will succeed and usually in one step. This answers a question raised near the end of subsection 30.3.4 about the possible instability of inverse iteration applied with an arbitrary starting vector and with the matrix $A - pI$, when p is accidentally close to an eigenvalue. We shall not get instability because we shall use several steps of inverse iteration, and rounding errors will guarantee that for some r successive vectors $\mathbf{y}^{(r)}$ in the iterative process are *not* orthogonal to the required eigenvector.

The only danger is with our triple-diagonal $C - \bar{\lambda}I$ in which $\bar{\lambda}$ is computed *separately* and we now seek the corresponding eigenvector. Our obvious method is implicitly equivalent to one step only of inverse iteration, and it is clear that what we have to do, to succeed in just one step, is to do this inverse iteration ***deliberately*** and ***properly***. In other words we solve the equations

$$(C - \bar{\lambda}I)\mathbf{x} = \mathbf{y}^{(0)},$$

with Gauss elimination with interchanges, and with a selected $\mathbf{y}^{(0)}$ which is not nearly orthogonal to the required eigenvector.

A good choice of $\mathbf{y}^{(0)}$ is not easy to guarantee absolutely, but the following method has rarely (say once in 10^4 problems) failed in practice. We reduce $C - \bar{\lambda}I$ to triangular form U with our stable method of elimination with interchanges, so that the equation left for solution is

$$U\mathbf{x} = \mathbf{c}^{(0)}.$$

Now we are not particularly interested in the precise values of the entries of $\mathbf{c}^{(0)}$, but merely that they are obtained from a truly arbitrary $\mathbf{y}^{(0)}$. So we now take

$$\mathbf{c}^{(0)T} = [1\ 1\ \ldots\ 1],$$

which corresponds to *some* particular $\mathbf{y}^{(0)}$, and it would be rather surprising if *this* $\mathbf{y}^{(0)}$ were nearly orthogonal to the required eigenvector. This, as we have said, almost always succeeds in one step, but for safety (and to eliminate the 1 in 10^4 failure rate), we can quite easily perform just one more step of inverse iteration and be virtually certain that the results merely confirm the accuracy of the first step!

Exercise

In Equation (1) of this sub-section we could work "backwards", taking $x_3 = 1$, computing x_2 from the last equation and then x_1 from the second. The first equation is then automatically satisfied if λ is an exact eigenvalue and we perform all subsequent arithmetic exactly. What happens if we use this method on the example of this sub-section and with which we failed catastrophically with the "forwards" method?

In fact, using precisely the same λ, working backwards and making no mistakes in the arithmetic we find a vector which, suitably scaled, agrees with the correct eigenvector to about nine figures! Can you explain this fact?

(N.B. This is an accident; the backwards method may fail in other examples, just as catastrophically as the forwards method failed in this example!)

Solution

In the backwards method all that happens in the previous theory is that $m\mathbf{e}_n$ is replaced everywhere by $m'\mathbf{e}_1$, where m' is some number and $\mathbf{e}_1$ is the first column of the unit matrix. Then, in Equation (6) of this sub-section

$$\mathbf{x} = \sum_{s=1}^{n} \frac{m'\mathbf{e}_1 \cdot \mathbf{x}_s}{\lambda_s - \bar{\lambda}} \mathbf{x}_s,$$

and $m'\mathbf{e}_1 \cdot \mathbf{x}_k$ *is* of reasonable size, since the first entry in $\mathbf{x}_k$ is 1. So our computed $\mathbf{x}$ is very nearly a multiple of the required eigenvector $\mathbf{x}_k$ whose eigenvalue λ_k is very close to $\bar{\lambda}$.

30.4.4 Error Analysis of the Similarity Transformation Method

For full analysis of induced instability we should consider:

(i) the similarity transformation from A to C;
(ii) computation of the eigenvalues of C;
(iii) the computation of the eigenvectors of C;
(iv) the "recovery" of the eigenvectors of A.

The analysis is rather long, so that we shall mention only the main points without going into the details.

In *Unit 8* we used a method of backward error analysis to show that the computed solution $\bar{\mathbf{x}}$ of $A\mathbf{x} = \mathbf{b}$ was the exact solution of

$$(A + \delta A)\bar{\mathbf{x}} = \mathbf{b} + \delta\mathbf{b},$$

where δA and $\delta\mathbf{b}$ have small entries. In the same way, we can show here for item (i) that the computed matrix $\hat{C}$, though possibly quite different from the exact matrix C computed (if this were possible) with exact arithmetic, is nevertheless the *exact* similarity transformation of *some* matrix $A + \delta A$, where the entries in the matrix δA are very small in relation to those of A.

There are various contributory factors to this result, but the most important is that the successive computed matrices $A_1, A_2, \ldots$ have entries which do not increase significantly in magnitude in relation to those of A. In *Unit 8* we related the instability of Gauss elimination *without* interchanges, for solving linear equations, to the growth of the successive matrices produced in the reduction of A to upper triangular U. The use of interchanges, giving multiples no greater than 1 in magnitude, limited this growth to such an extent that we maintained good stability, and the *computed* U was the U which would be obtained by *exact* arithmetic from an original $A + \delta A$ with small δA.

It is the choice of our particularly simple *orthogonal* matrices P which produces the analogous effect in our transformation of A to the triple-diagonal C.

Item (iv), the recovery of the eigenvectors of A, turns out to be stable for quite similar reasons, since we obtain the eigenvectors of A by premultiplying those of C by a product of these simple orthogonal matrices of type P.

For item (ii), the computation of the eigenvalues of C, we can show that the use of the recurrence relation for computing

$$f_{r+1} = (a_{r+1} - p)f_r - b_{r+1}^2 f_{r-1}$$

actually produces an f_{r+1} which satisfies exactly

$$f_{r+1} = (a'_{r+1} - p)f_r - (b'_{r+1})^2 f_{r-1},$$

where a'_{r+1} and b'_{r+1} differ by very small amounts from a_{r+1} and b_{r+1} respectively. This means that we are finding the exact eigenvalues of a matrix differing only slightly from the original triple-diagonal matrix C.

In Items (i), (ii) and (iv), then, there is virtually no induced instability; the eigenvalues are always produced with great accuracy, and the eigenvectors of A have little more error than those of C. It is only in item (iii), the computation of the eigenvectors of C, that we have to be careful. We have already seen in the preceding sub-section how one obvious method of calculating the eigenvectors may suffer from induced instability, but we have also seen how this instability can be avoided by using inverse iteration deliberately and properly. When we use this method of finding the eigenvectors of C the entire calculation is therefore free from induced instability, the *computed* results being the *exact* results for a slightly different matrix $A + \delta A$ with a "small" δA. Our work on inherent instability then guarantees that our computed eigenvalues are very accurate, and the eigenvectors will be as good as we can expect in virtue of the possible ill-conditioning of eigenvectors corresponding to nearly equal eigenvalues.

30.4.5 Summary of Section 30.4

In this section we defined the terms

similarity transformation	(page C31)	*
triple-diagonal matrix	(page C31)	*
leading or principal submatrix	(page C35)	*
separation of eigenvalues	(page C35)	*
bisection process	(page C36)	*

We introduced the notation

P_r	(page C31)
A_r	(page C31)
s_r, c_r	(page C32)
C	(page C33)
f_r	(page C35)

Theorems

1. (page C35)
For a symmetric matrix, the eigenvalues of the principal submatrix of order r separate those of the principal submatrix of order $r + 1$. *

2. (page C35)
If the determinants of successive principal submatrices of the matrix $A - pI$ are $f_1, f_2, \ldots, f_n$, with $f_0 = 1$, then the number of eigenvalues of A which are greater than p is equal to the number of agreements in sign in successive members of the sequence $f_0, f_1, \ldots, f_n$. *

Techniques

1. Find orthogonal plane-rotation matrices to reduce to zero particular entries of a matrix by a similarity transformation. *
2. Transform a symmetric matrix to symmetric triple-diagonal form. *
3. Compute principal determinants of a triple-diagonal matrix using a recurrence relation, and find the number of agreements in sign in successive members of the sequence. *
4. Use this result to determine how many eigenvalues are greater than some number, and then to bracket a required eigenvalue in intervals of decreasing size. *
5. Avoid the induced instability of an obvious method for computing eigenvectors by using one or two steps of inverse iteration instead. *

Formula

The principal determinants $f_1, f_2, f_3, \ldots,$ of a triple-diagonal matrix

$$\begin{bmatrix} a_1 - p & b_2 & 0 & \cdots & \\ b_2 & a_2 - p & b_2 & 0 & \cdots \\ 0 & b_3 & a_3 - p & \ddots & \\ \vdots & 0 & & & \\ \vdots & \vdots & & & \end{bmatrix}$$

satisfy

$$f_{r+1} = (a_{r+1} - p)f_r - b_{r+1}^2 f_{r-1}$$

30.5 ERROR AND CORRECTION OF APPROXIMATE SOLUTIONS (OPTIONAL)

30.5.0 Introduction

The method of backward error analysis, as we know, measures the stability of our method but not the accuracy of the computed solution. In this section we shall see how to assess the error of an approximate eigensolution, and to improve the eigenvalue, with a relatively small amount of extra work.

In *Unit 8* we looked at the similar problem relevant to algebraic equations $A\mathbf{x} = \mathbf{b}$, and we observed that the size of the residual vector $\mathbf{r} = \mathbf{b} - A\bar{\mathbf{x}}$, for an approximate solution $\bar{\mathbf{x}}$, gave little or no information about the error $\bar{\mathbf{x}} - \mathbf{x}$ of this approximation. For the symmetric eigenvalue problem, however, we shall see that the residual vector turns out to be both useful and informative.

30.5.1 Error of an Approximate Eigensolution

Suppose that λ and $\mathbf{x}$ are approximations to an eigensolution $\lambda_1, \mathbf{x}_1$. Assuming $\mathbf{x}$ to be normalized, we compute the residual vector

$$\mathbf{r} = A\mathbf{x} - \lambda\mathbf{x}.$$

Then, if we define ε by

$$\varepsilon = \sqrt{(\mathbf{r} \cdot \mathbf{r})},$$

we shall prove that

$$|\lambda_1 - \lambda| \leqslant \varepsilon \tag{1}$$

To prove this we write, as usual,

$$\mathbf{x} = \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n \tag{2}$$

where the eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are normalized. Then since $\mathbf{x}_i \cdot \mathbf{x}_j = \delta_{ij}$, we have

$$1 = \mathbf{x} \cdot \mathbf{x} = \alpha_1^2 + \cdots + \alpha_n^2 \tag{3}$$

The corresponding expression for $\mathbf{r}$ is

$$\mathbf{r} = (A - \lambda I)\mathbf{x} = (\lambda_1 - \lambda)\alpha_1\mathbf{x}_1 + \cdots + (\lambda_n - \lambda)\alpha_n\mathbf{x}_n$$

from which we obtain

$$\begin{aligned} \varepsilon^2 = \mathbf{r} \cdot \mathbf{r} &= (\lambda_1 - \lambda)^2\alpha_1^2 + \cdots + (\lambda_n - \lambda)^2\alpha_n^2 \\ &\geqslant (\lambda_1 - \lambda)^2(\alpha_1^2 + \cdots + \alpha_n^2) \end{aligned} \tag{4}$$

assuming that λ is closer to λ_1 than to any other eigenvalue. Using (3) and rearranging, we obtain the result (1) which we wanted to prove.

This is clearly a very useful and easily computed upper bound for the error of our computed eigenvalue, and we notice that it does not depend on the properties of any other eigensolution. We shall find that the corresponding result for the eigenvector depends on the nearness of another eigenvalue to λ_1, and this is hardly surprising when we recall the possibilities of inherent instability of an eigenvector.

To estimate the error $\mathbf{x} - \mathbf{x}_1$ in the approximate eigenvector $\mathbf{x}$, we use (2) and write

$$\mathbf{x} - \mathbf{x}_1 = (\alpha_1 - 1)\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_n\mathbf{x}_n,$$

so that

$$(\mathbf{x}-\mathbf{x}_1)\cdot(\mathbf{x}-\mathbf{x}_1) = (\alpha_1-1)^2+\alpha_2^2+\cdots+\alpha_n^2$$
$$= 2(1-\alpha_1) \qquad (5)$$

by Equation (3).

Our problem is therefore reduced to an estimation of α_1. For this purpose we define a to be the difference between the approximate eigenvalue λ and the *nearest* eigenvalue other than λ_1, so that

$$|\lambda_s-\lambda| \geqslant a, \qquad s=2,3,\ldots,n.$$

Then (4) gives

$$\begin{aligned}\varepsilon^2 &= (\lambda_1-\lambda)^2\alpha_1^2+(\lambda_2-\lambda)^2\alpha_2^2+\cdots+(\lambda_n-\lambda)^2\alpha_n^2\\ &\geqslant a^2(\alpha_2^2+\cdots+\alpha_n^2)\\ &= a^2(1-\alpha_1^2).\end{aligned} \qquad (6)$$

So

$$\alpha_1^2 \geqslant 1-\frac{\varepsilon^2}{a^2},$$

and if we neglect fourth and higher powers of ε and assume, reasonably enough, that α_1 is positive, this gives simply

$$\alpha_1 \geqslant 1-\frac{1}{2}\frac{\varepsilon^2}{a^2}.$$

Hence

$$2(1-\alpha_1) \leqslant \frac{\varepsilon^2}{a^2},$$

and (5) finally gives the required result

$$(\mathbf{x}-\mathbf{x}_1)\cdot(\mathbf{x}-\mathbf{x}_1) \leqslant \frac{\varepsilon^2}{a^2}. \qquad (7)$$

This says that the sum of the squares of the entries in $\mathbf{x}-\mathbf{x}_1$ is at most $\frac{\varepsilon^2}{a^2}$, and the individual entries, the errors in those of $\mathbf{x}$ in relation to those of $\mathbf{x}_1$, can hardly exceed $\frac{\varepsilon}{a}$ in magnitude.

Example

In the Example of sub-section 30.3.2 we found the approximations

$$\mathbf{y} = [0.731 \quad 0.233 \quad 1.000]^T, \lambda = 6.43$$

for the largest eigensolution of the matrix

$$A = \begin{bmatrix} 2 & 1 & 3\\ 1 & -1 & 1\\ 3 & 1 & 4\end{bmatrix}.$$

Let us find upper bounds for the errors of these results.

The required normalized approximation $\mathbf{x}$ is $\mathbf{y}/k$, where $k=(\mathbf{y}\cdot\mathbf{y})^{1/2}$, but the arithmetic is simplified (and more accurate!) if we defer the normalization.

Computation gives

$$k^2 = \mathbf{y}\cdot\mathbf{y} = 1.588\,650,$$

and

$$\begin{aligned}\mathbf{r} &= (A-\lambda I)\mathbf{x} = (A-\lambda I)\mathbf{y}/k\\ &= \frac{1}{k}[-0.005\,33 \quad -0.000\,19 \quad -0.004\,00]^T\end{aligned}$$

so that

$$\varepsilon^2 = \mathbf{r}\cdot\mathbf{r} = \frac{(0.005\,33)^2 + (0.000\,19)^2 + (0.004\,00)^2}{k^2}$$

$$= \frac{0.000\,044\,445}{1.588\,650}$$

$$= 0.000\,027\,98$$

to four significant figures, giving

$$\varepsilon \simeq 0.0053.$$

By Equation (1) the error in the computed eigenvalue is then less than 0.0053 in magnitude. Moreover, since the eigenvalues of A are approximately 6.4, -1.3 and -0.1, the next nearest eigenvalue is at a distance of about 6.5 from λ_1, so that $a \simeq 6.5$ and the error in $\mathbf{x}$, the normalized approximate eigenvector, cannot (by (7)) exceed $\varepsilon/a \simeq 0.0008$ in any entry.

Accurate computation shows that the error in the eigenvalue is 0.0050, and the maximum error in the eigenvector is 0.0003 in its first entry. Our error *bounds*, 0.0053 and 0.0008, do not therefore greatly overestimate the true errors.

Exercise

Repeat the process of the example with the approximation

$$\mathbf{y} = [0.7 \quad 0.2 \quad 1.0]^T, \quad \lambda = 6.$$

Solution

We find

$$(A - 6I)\mathbf{y} = [0.4 \quad 0.3 \quad 0.3]^T$$

and

$$\mathbf{y}\cdot\mathbf{y} = 1.53,$$

so that

$$\varepsilon^2 = \frac{0.34}{1.53} = 0.222$$

and

$$\varepsilon = 0.47, \text{ approximately.}$$

So the error in the eigenvalue is at most 0.47, and in the normalized eigenvector at most 0.08 in any component.

30.5.2 Correction of an Approximate Eigensolution

Finally, we examine the possibility of improving our approximation with little extra work. This we can do very easily for the eigenvalue, the more accurate estimate corresponding to our approximate eigenvector $\mathbf{y}$ being the *Rayleigh quotient*

$$\lambda_R = \frac{\mathbf{y}^T A\mathbf{y}}{\mathbf{y}\cdot\mathbf{y}} = \mathbf{x}^T A\mathbf{x}$$

where

$$\mathbf{x} = \frac{\mathbf{y}}{(\mathbf{y}\cdot\mathbf{y})^{1/2}}$$

is the corresponding normalized approximate eigenvector. The reason why the Rayleigh quotient is a good estimate can be seen by writing

$$\mathbf{x} = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_n \mathbf{x}_n$$

so that

$$\lambda_R = \alpha_1^2 \lambda_1 + \alpha_2^2 \lambda_2 + \cdots + \alpha_n^2 \lambda_n \tag{1}$$

by Formula (5) of sub-section 30.1.1. It follows, since

$$\alpha_1^2 + \cdots + \alpha_n^2 = 1,$$

that

$$\begin{aligned}\lambda_R - \lambda_1 &= (\alpha_1^2 \lambda_1 + \cdots + \alpha_n^2 \lambda_n) - (\alpha_1^2 + \cdots + \alpha_n^2)\lambda_1 \\ &= \alpha_2^2(\lambda_2 - \lambda_1) + \cdots + \alpha_n^2(\lambda_n - \lambda_1)\end{aligned}$$

It follows that if α denotes the error in $\mathbf{x}$ as an approximation to $\mathbf{x}_1$, measured by the largest of the numbers $|\alpha_2|, \ldots, |\alpha_n|$, then the error $\lambda_R - \lambda_1$ in the improved approximate eigenvalue λ_R is roughly α^2, and is likely to be much smaller than the error of the approximate eigenvector it was obtained from.

We can even find a bound for the error $|\lambda_R - \lambda_1|$ of the Rayleigh estimate. Using (1) and the normalization condition

$$\alpha_1^2 + \cdots + \alpha_n^2 = 1,$$

we can write

$$\lambda_R(\alpha_1^2 + \cdots + \alpha_n^2) = \lambda_1 \alpha_1^2 + \cdots + \lambda_n \alpha_n^2$$

so that

$$\begin{aligned}(\lambda_R - \lambda_1)\alpha_1^2 &= (\lambda_2 - \lambda_R)\alpha_2^2 + \cdots + (\lambda_n - \lambda_R)\alpha_n^2 \\ &= \frac{(\lambda_2 - \lambda_R)^2 \alpha_2^2}{\lambda_2 - \lambda_R} + \cdots + \frac{(\lambda_n - \lambda_R)^2 \alpha_n^2}{\lambda_n - \lambda_R}\end{aligned}$$

It follows that

$$\begin{aligned}|\lambda_R - \lambda_1|\alpha_1^2 &\leqslant \frac{(\lambda_2 - \lambda_R)^2 \alpha_2^2}{|\lambda_2 - \lambda_R|} + \cdots + \frac{(\lambda_n - \lambda_R)^2 \alpha_n^2}{|\lambda_n - \lambda_R|} \\ &\leqslant \frac{1}{a}[(\lambda_2 - \lambda_R)^2 \alpha_2^2 + \cdots + (\lambda_n - \lambda_R)^2 \alpha_n^2]\end{aligned}$$

where a is the smallest of the denominators. Defining

$$\begin{aligned}\mathbf{r} &= (A - \lambda_R)\mathbf{x} \\ &= (\lambda_1 - \lambda_R)\alpha_1 \mathbf{x}_1 + \cdots + (\lambda_n - \lambda_R)\alpha_n \mathbf{x}_n,\end{aligned}$$

we obtain

$$|\lambda_R - \lambda_1|\alpha_1^2 \leqslant \frac{1}{a}\mathbf{r} \cdot \mathbf{r} = \frac{1}{a}\varepsilon^2$$

so that

$$\begin{aligned}|\lambda_R - \lambda_1| &\leqslant \frac{\varepsilon^2}{a\alpha_1^2} \\ &\leqslant \frac{\varepsilon^2}{a(1 - \varepsilon^2/a^2)} \text{ by (6) of sub-section 30.5.1.}\end{aligned}$$

A reasonable knowledge of a thus gives a very useful estimate, approximately ε^2/a, for the error of the Rayleigh quotient.

Unfortunately we have no simple method for correcting or improving our approximate vector. We were able to improve the eigenvalue by evaluating the Rayleigh estimate because the error in our approximation did not depend on any other eigensolution; but we have no simple methods for correcting the Rayleigh estimate, or of obtaining any simple correction to the approximate eigenvector, because the error bounds for these approximations do need a knowledge of the other eigensolutions. Knowing approximations to all the eigensolutions we can in fact correct the approximate eigenvectors but this is a much more elaborate operation which we shall not consider here.

Example

Let us find (i) the Rayleigh estimate for the approximate eigenvector

$$\mathbf{y} = [0.731 \quad 0.233 \quad 1.000]^T$$

for the matrix

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 3 & 1 & 4 \end{bmatrix}$$

used in the Example of sub-section 30.5.1, and (ii) a bound for its error.

We find

$$A\mathbf{y} = [4.695 \quad 1.498 \quad 6.426]^T$$

and with

$$\mathbf{y}^T\mathbf{y} = 1.588\,650$$

we get

$$\lambda_R = \frac{\mathbf{y}^T A\mathbf{y}}{\mathbf{y}^T\mathbf{y}} = \frac{10.207\,079}{1.588\,650} = 6.425\,001\,7 \text{ approximately.}$$

For the residual vector we find

$$\begin{aligned}\mathbf{r} &= (\mathbf{y}\cdot\mathbf{y})^{-1/2}(A\mathbf{y} - \lambda_R\,\mathbf{y}) \\ &= (\mathbf{y}\cdot\mathbf{y})^{-1/2}[-0.001\,68 \quad 0.000\,97 \quad 0.001\,00]^T\end{aligned}$$

(to five decimals), so that

$$\varepsilon^2 = 0.000\,003\,00, \frac{\varepsilon^2}{a} \simeq \frac{\varepsilon^2}{6.5} \simeq 0.000\,000\,5.$$

This is an error bound for the Rayleigh quotient, its actual error being about 0.000 000 4. We note the great improvement, in our estimate of the eigenvalue, which we obtain by computing the Rayleigh quotient.

Exercises

1. If $\mathbf{y}$ is an approximate eigenvector, λ an approximate eigenvalue, and $\mathbf{r} = A\mathbf{y} - \lambda\mathbf{y}$, show that the Rayleigh quotient gives a correction to λ of amount $\dfrac{\mathbf{y}^T\mathbf{r}}{\mathbf{y}^T\mathbf{y}}$.

2. Using the method of Exercise 1, and the result of the exercise of the previous sub-section, show that, for the vector

$$\mathbf{y} = [0.7 \quad 0.2 \quad 1.0]^T,$$

the Rayleigh quotient gives $\lambda_R = 6.42$, approximately.

Solutions

1. Since

$$A\mathbf{y} - \lambda\mathbf{y} = \mathbf{r},$$

and

$$\lambda_R = \frac{\mathbf{y}^T A\mathbf{y}}{\mathbf{y}^T\mathbf{y}}$$

we have

$$\lambda_R = \frac{\mathbf{y}^T(\lambda\mathbf{y} + \mathbf{r})}{\mathbf{y}^T\mathbf{y}} = \lambda + \frac{\mathbf{y}^T\mathbf{r}}{\mathbf{y}^T\mathbf{y}}.$$

2. The exercise of the previous sub-section gives

$$\mathbf{r} = [0.4 \quad 0.3 \quad 0.3]^T, \mathbf{y}^T\mathbf{y} = 1.53, \text{ and } \lambda = 6,$$

so that

$$\lambda_R = 6 + \left(\frac{(0.4)(0.7) + (0.3)(0.2) + (0.3)(1.0)}{1.53}\right) = 6.42,$$

to two decimal places. This has a maximum error of little more than 0.005, although the vector is correct to only one significant figure.

30.5.3 Summary of Section 30.5

In this section we defined the term

Rayleigh quotient (page C45)

and introduced the notation

λ_R (page C45)

Techniques

1. Calculate a bound for the error of an approximate eigenvalue by computing the residual vector of an approximate eigensolution.
2. With a knowledge of the next nearest eigenvalue, calculate a bound for the error of the approximate eigenvector.
3. Improve the approximate eigenvalue by calculating the Rayleigh quotient, and determine an upper bound for the error of this approximation to the eigenvalue.

30.6 SUMMARY OF THE UNIT

In the first section we established some properties of eigenvalues and eigenvectors, for symmetric matrices, needed in the subsequent numerical work. These included measures for the "magnitudes" of matrices and vectors. We also observed, and illustrated, the lack of economy and stability in the use of the characteristic equation for computing eigenvalues.

The second section discussed the problem of inherent instability, and we showed that the eigenvalues are always "well-conditioned" in relation to small perturbations in the entries of the matrix, and that the eigenvectors are ill-conditioned only when associated with eigenvalues which are nearly equal.

In the third section we discussed a method of *direct* iteration with the matrix $A - pI$, choosing p to give convergence to one or other of the extreme eigensolutions, the only possibilities with this method. Then we showed that *inverse* iteration with $A - pI$ (direct iteration with $(A - pI)^{-1}$) will converge to the eigensolution whose eigenvalue is nearest to p, so that by suitable choice of p we can find *any* eigensolution. The iterative methods were expressed in practical computational form, and we illustrated the method of Gauss elimination with interchanges for solving the linear equations involved in inverse iteration. We also showed that all these methods are virtually free from induced instability.

In the fourth section we showed how we could transform the given matrix in a finite sequence of orthogonal similarity transformations into a simpler form, the triple-diagonal form which is very suitable for computation. We showed how to compute determinants of successive principal submatrices of $C - pI$ by using a simple recurrence relation, and observed that by inspecting the signs of members of this sequence we could find the number of eigenvalues exceeding the number p. A systematic bisection process then located an eigenvalue quickly and accurately. The obvious method of computation of the eigenvector, however, we found to exhibit significant induced instability, and replaced this by inverse iteration which always converged in one or at most two steps. With this refinement all our methods were free from induced instability.

In the final (optional) section we found bounds for the error in an approximate eigensolution, and showed how the computation of the Rayleigh quotient, an economic operation, gave a much better estimate for the eigenvalue.

Definitions

eigensystem	(page C5)	* *
modal matrix	(page C7)	*
scaling of matrices	(page C9)	* *
perturbation of a matrix	(page C13)	* *
first-order perturbation of an eigenvalue	(page C13)	* *
first-order perturbation of an eigenvector	(page C13)	* *
direct iteration with A and with $A - pI$	(page C20)	* *
inverse iteration with A and with $A - pI$	(page C27)	* *
similarity transformation	(page C31)	*
triple-diagonal matrix	(page C31)	*
leading or principal submatrix	(page C35)	*
separation of eigenvalues	(page C35)	*
bisection process	(page C36)	*
Rayleigh quotient	(page C45)	

Theorems

1. (page C35)

For a symmetric matrix, the eigenvalues of the principal submatrix of order r separate those of the principal submatrix of order $r + 1$. *

2. (page C35)

If the determinants of successive principal submatrices of the matrix $A - pI$ are $f_1, f_2, \ldots, f_n$, with $f_0 = 1$, then the number of eigenvalues of A which are greater than p is equal to the number of agreements in sign in successive members of the sequence $f_0, f_1, \ldots, f_n$. *

Techniques

1. Scale a matrix so that the magnitude of the largest entry of the new matrix is 1. * *
2. Use direct iteration, with the matrix $A - pI$, to converge to one of the extreme eigensolutions. *
3. Find a p which gives most rapid convergence to one of these eigensolutions. *
4. Use inverse iteration with A to converge to the eigensolution corresponding to the eigenvalue of smallest magnitude. * *
5. Use inverse iteration with $A - pI$ to converge to the eigensolution whose eigenvalue is closest to a given p. * *
6. Use stable Gauss elimination with interchanges for solving linear equations in inverse iteration. *
7. Find orthogonal plane-rotation matrices to reduce to zero particular entries of a matrix by a similarity transformation. *
8. Transform a symmetric matrix to symmetric triple-diagonal form. *
9. Compute principal determinants of a triple-diagonal matrix using a recurrence relation, and find the number of agreements in sign in successive members of the sequence. *
10. Use this result to determine how many eigenvalues are greater than some number, and then to bracket a required eigenvalue in intervals of decreasing size. *
11. Avoid the induced instability of an obvious method for computing eigenvectors by using one or two steps of inverse iteration instead. *
12. Calculate a bound for the error of an approximate eigenvalue by computing the residual vector of an approximate eigensolution.
13. With a knowledge of the next nearest eigenvalue, calculate a bound for the error of the approximate eigenvector.
14. Improve the approximate eigenvalue by calculating the Rayleigh quotient, and determine an upper bound for the error of this approximation to the eigenvalue.

Notation

$\lambda_{\max}$	(page C8)
$M(\mathbf{x})$	(page C8)
$M(A)$	(page C8)
$\lambda_r(\varepsilon)$	(page C13)
$\mathbf{x}_r(\varepsilon)$	(page C13)
$m(\mathbf{z})$	(page C22)
P_r	(page C31)
A_r	(page C31)
s_r, c_r	(page C32)
C	(page C33)
f_r	(page C35)
λ_R	(page C45)

30.7 SELF-ASSESSMENT

Self-assessment Test

This Self-assessment Test is designed to help you test your understanding of the unit. It can also be used, together with the summary of the unit, for revision. The answers to these questions will be found on the next non-facing page. We suggest that you complete the whole test before looking at the answers.

(In these questions all the matrices are symmetric.)

1. (i) For any vector y we know that

$$\mathbf{y}^T A\mathbf{y} \leqslant \mathbf{y}^T\mathbf{y}\,|\lambda_{\max}|,$$

where $\lambda_{\max}$ is the eigenvalue of largest magnitude of a matrix A. For what vector y do we have equality in this expression?

(ii) (optional)
If y is near to some particular eigenvector of A, what can we say about the number $\dfrac{\mathbf{y}^T A\mathbf{y}}{\mathbf{y}^T\mathbf{y}}$?

(iii) Define the number $M(A)$ for a matrix A, and say what value $M(A)$ has for the matrix

$$A = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}.$$

Verify that

$$M(A) \geqslant |\lambda_i| \qquad (i = 1, 2)$$

where λ_1 and λ_2 are the eigenvalues of A.

2. (i) A matrix A_1 has eigenvalues 3.1, -0.6 and -0.8, and a matrix A_2 *with the same eigenvectors* as A_1, has eigenvalues 2.9, -3.8, and -4.8. For which of these matrices are we more likely to have significant ill-conditioning in (a) the eigenvalues, and (b) the eigenvectors, in respect of the same small perturbations in the entries of the matrices?

(ii) Which of the eigenvectors are likely to exhibit the greater degree of ill-conditioning in 2(i), and what is the approximate ratio of the magnitudes of this ill-conditioning for the matrices A_1 and A_2?

3. We propose to compute the eigensolutions of the matrices A_1 and A_2 in Question 2(i) by an iterative method.

(i) To what eigenvalues shall we converge using direct iteration with $A_1 - pI$ and $A_2 - pI$ for

(a) $p = 0$,
(b) $p = -1$,
(c) $p = 50$?

(ii) What values of p would we use to obtain most rapid convergence to

(a) the eigenvalue 3.1 of A_1,
(b) the eigenvalue -4.8 of A_2,
(c) the eigenvalue -3.8 of A_2?

(iii) If we use inverse iteration with $A_1 - pI$ and $A_2 - pI$, to what eigenvalues shall we converge with $p = 1$, and how can we estimate the rate of convergence?

4. (i) Reduce to triple-diagonal form C, by means of an orthogonal similarity transformation, the matrix

$$A = \begin{bmatrix} 1 & \sqrt{2} & \sqrt{2} \\ \sqrt{2} & -\sqrt{2} & -1 \\ \sqrt{2} & -1 & \sqrt{2} \end{bmatrix}$$

(ii) An eigenvector of C is

$$\mathbf{x}_1 = \left[-\frac{1}{\sqrt{2}} \quad 0 \quad 1\right]^T.$$

Write down the corresponding *normalized* eigenvector of A.

(iii) Find the number of positive eigenvalues of A.

5. The eigenvalues of the triple-diagonal matrix

$$C = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & \sqrt{2} \\ 0 & \sqrt{2} & 1 \end{bmatrix}$$

are arranged in order $\lambda_1 > \lambda_2 > \lambda_3$.

(i) All the eigenvalues lie in the interval $(-a, a)$. What is the smallest value of a that you can obtain almost by inspection from the matrix C?

(ii) By taking $a = 6$, for simplicity, find an interval which contains only the eigenvalue λ_2.

6. (Optional)
For the matrix of Question 5 we have computed the approximate eigensolution (with normalized eigenvector)

$$\bar{\lambda} = 2, \quad \mathbf{x}^T = \left[\frac{1}{\sqrt{2}} \quad \frac{1}{2} \quad \frac{1}{2}\right].$$

(i) Show that the error of this approximate eigenvalue does not exceed $(9 - 6\sqrt{2})^{1/2}$. (This is just larger than 0.7, and the correct eigenvalue is $\sqrt{7} \simeq 2.65$.)

(ii) Find the Rayleigh quotient for this vector, and observe its great accuracy as an approximation to the eigenvalue (the approximate eigenvector has only the first figure accurate in all three components).

Solutions to Self-assessment Test

1. (i) The required vector is any multiple of the eigenvector corresponding to the eigenvalue λ_{max}.

(ii) The number $y^T Ay/y^T y$ is the Rayleigh quotient. If y is close to some eigenvector then the Rayleigh quotient is even closer to the corresponding eigenvalue.

(iii) $M(A)$ is the maximum row sum of absolute values of the entries of A. For the given matrix $M(A) = 6$. The eigenvalues satisfy

$$(1 - \lambda)(4 - \lambda) - 2^2 = 0, \text{ or } \lambda^2 - 5\lambda = 0.$$

Then $\lambda_1 = 0$, $\lambda_2 = 5$, and

$$M(A) > |\lambda_i| \qquad (i = 1, 2)$$

2. (i) (a) Neither matrix exhibits significant ill-conditioning in the eigenvalues, and the "first-order" perturbations in the eigenvalues corresponding to the same eigenvector have the same value for each matrix.

(b) Two of the eigenvectors of A_1 are more likely to be ill-conditioned than any eigenvectors of A_2, in virtue of the near equality of the eigenvalues -0.6 and -0.8 of A_1. The eigenvalues of A_2 are well separated and its eigenvectors therefore well-conditioned.

(ii) Each of the eigenvectors corresponding to the eigenvalues -0.6 and -0.8 of A_1 will exhibit ill-conditioning, by a factor proportional to $1/(-0.6 + 0.8) = 5$. The corresponding number for the matrix A_2 is 1, so the relevant eigenvectors of A_1 are likely to be "5 times as ill-conditioned" as the corresponding eigenvectors of A_2.

3. (i) (a) With $p = 0$, we converge with direct iteration with A_1 to the eigenvalue 3.1, and with A_2 to the eigenvalue -4.8, those farthest from the origin.

(b) With $p = -1$, we converge with direct iteration with $A_1 - pI$ to the eigenvalue 3.1 and with $A_2 - pI$ to the eigenvalue 2.9, those for which $|\lambda_r - p|$ is largest.

(c) With $p = 50$, we get convergence, for the same reason, to the eigenvalues -0.8 and -4.8.

(ii) (a) and (b) The required values of p are

$$\tfrac{1}{2}(-0.6 - 0.8) = -0.7, \text{ and } \tfrac{1}{2}(2.9 - 3.8) = -0.45.$$

(c) No value of p will give convergence to the eigenvalue -3.8, the "non-extreme" eigenvalue of A_2.

(iii) With $p = 1$ we converge to the eigenvalue for which $|\lambda_r - p|$ is smallest, that is the eigenvalues -0.6 of A_1 and 2.9 of A_2. The rates of convergence depend on the ratios of

$$|(\lambda_k - p)/(\lambda_r - p)|,$$

where λ_k is nearest to p, and λ_r is another eigenvalue. For the first result of (iii) we converge to the eigenvalue -0.6 at a rate at which $\left(\frac{1.6}{1.8}\right)^s$ and $\left(\frac{1.6}{2.1}\right)^s$ get small as s increases and for the second result the corresponding numbers are $\left(\frac{1.9}{4.8}\right)^s$ and $\left(\frac{1.9}{5.8}\right)^s$.

4. (i) The required orthogonal matrix is a plane-rotation in the (2, 3) plane through angle θ, where

$$\cos\theta = a_{21}/(a_{21}^2 + a_{31}^2)^{1/2},$$
$$\sin\theta = a_{31}/(a_{21}^2 + a_{31}^2)^{1/2},$$

giving $\cos\theta = \sin\theta = \dfrac{1}{\sqrt{2}}$.

Then $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \\ 0 & \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{bmatrix}$.

and by computation we find that the required triple-diagonal form (only one rotation being needed) is

$$P^TAP = C = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & \sqrt{2} \\ 0 & \sqrt{2} & 1 \end{bmatrix}$$

(ii) The eigenvector of A corresponding to an eigenvector $\mathbf{x}_1$ of C is $P\mathbf{x}_1$, which here is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \dfrac{1}{\sqrt{2}} & \dfrac{-1}{\sqrt{2}} \\ 0 & \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -\dfrac{1}{\sqrt{2}} \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\dfrac{1}{\sqrt{2}} \\ -\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix}$$

This eigenvector $\mathbf{x}$, normalized so that $\mathbf{x}^T\mathbf{x} = 1$, is therefore

$$\mathbf{x}^T = \frac{1}{\sqrt{3}} [1 \quad 1 \quad -1].$$

(iii) The number of positive eigenvalues of A is the same as the corresponding number for C. Using the relevant recurrence relation (or by direct calculation) we find the sequence

f_0	f_1	f_2	f_3
1	1	-5	-7

for successive determinants f_1, f_2 and f_3 (with $f_0 = 1$) of principal submatrices of $C - 0I = C$. There are two agreements in sign and therefore two positive eigenvalues (greater than $p = 0$).

5. (i) No eigenvalue exceeds $M(C)$ in magnitude, so that $a = M(C)$ is the smallest easily calculable value of a. We find $M(C) = 3 + \sqrt{2}$ (from the second row).

(ii) Taking $p = 0$ (bisecting the interval $(-6, 6)$) we find for successive principal submatrices of C the sequence

f_0	f_1	f_2	f_3
1	1	-5	-7

with two agreements in sign. Then there are two eigenvalues in the interval (0, 6), which must be λ_1 and λ_2. We therefore take $p = 3$, and find for $C - 3I$ the sequence

f_0	f_1	f_2	f_3
1	-2	4	-4

There are no agreements in sign, so that both positive roots are in the interval (0, 3].

Now take $p = 1.5$, and for $C - 1.5I$ we find

f_0	f_1	f_2	f_3
1	−0.5	−2.75	2.375

There is now one agreement in sign, so that λ_1 is in the interval (1.5, 3], and λ_2 in the interval (0, 1.5].

6. (i) Since the approximate eigenvalue is normalized, the error bound for the approximate eigenvalue $\bar{\lambda}$ is $(\mathbf{r} \cdot \mathbf{r})^{1/2}$, where $\mathbf{r} = (C - \bar{\lambda}I)\mathbf{x}$. We find

$$\mathbf{r}^T = [1 - \tfrac{1}{2}\sqrt{2} \quad -\tfrac{3}{2} + \tfrac{3}{2}\sqrt{2} \quad -\tfrac{1}{2} + \tfrac{1}{2}\sqrt{2}],$$

and a little computation gives $\mathbf{r} \cdot \mathbf{r} = 9 - 6\sqrt{2}$. The error bound for $\bar{\lambda}$ is therefore $(9 - 6\sqrt{2})^{1/2}$.

(ii) Since the approximate eigenvector is normalized, the Rayleigh quotient is just $\mathbf{x}^T C\mathbf{x}$, which is equal to $\bar{\lambda} + \mathbf{x}^T\mathbf{r}$. We find

$$\mathbf{x}^T\mathbf{r} = \tfrac{3}{2}(\sqrt{2} - 1) \simeq 0.62,$$

so that the Rayleigh quotient is approximately 2.62, agreeing to *two* figures with the correct eigenvalue.

LINEAR MATHEMATICS

1 Vector Spaces
2 Linear Transformations
3 Hermite Normal Form
4 Differential Equations I
5 Determinants and Eigenvalues
6 NO TEXT
7 Introduction to Numerical Mathematics: Recurrence Relations
8 Numerical Solution of Simultaneous Algebraic Equations
9 Differential Equations II: Homogeneous Equations
10 Jordan Normal Form
11 Differential Equations III: Nonhomogeneous Equations
12 Linear Functionals and Duality
13 Systems of Differential Equations
14 Bilinear and Quadratic Forms
15 Affine Geometry and Convex Cones
16 Euclidean Spaces I: Inner Products
17 NO TEXT
18 Linear Programming
19 Least-squares Approximation
20 Euclidean Spaces II: Convergence and Bases
21 Numerical Solution of Differential Equations
22 Fourier Series
23 The Wave Equation
24 Orthogonal and Symmetric Transformations
25 Boundary-value Problems
26 NO TEXT
27 Chebyshev Approximation
28 Theory of Games
29 Laplace Transforms
30 Numerical Solution of Eigenvalue Problems
31 Fourier Transforms
32 The Heat Conduction Equation
33 Existence and Uniqueness Theorem for Differential Equations
34 NO TEXT

335 01120 9